

1 ROB BONTA, SBN 202668
Attorney General of California
2 ANTHONY R. HAKL, SBN 197335
Supervising Deputy Attorney General
3 ANNA FERRARI, SBN 261579
GABRIELLE D. BOUTIN, SBN 267308
4 Deputy Attorneys General
1300 I Street, Suite 125
5 P.O. Box 944255
Sacramento, CA 94244-2550
6 Telephone: (916) 210-6053
Fax: (916) 324-8835
7 E-mail: Gabrielle.Boutin@doj.ca.gov
Attorneys for Defendant Attorney
8 General Rob Bonta, in his official
capacity

9
10 IN THE UNITED STATES DISTRICT COURT
11 FOR THE EASTERN DISTRICT OF CALIFORNIA
12 SACRAMENTO DIVISION
13

14 **X CORP. ,**

15 Plaintiff,

16 **v.**

17 **ROBERT A. BONTA, ATTORNEY GENERAL OF**
18 **CALIFORNIA, IN HIS OFFICIAL CAPACITY,**

19 Defendant.
20
21
22

2:23-CV-01939-WBS-AC

**DECLARATION OF GABRIELLE
BOUTIN IN SUPPORT OF
DEFENDANT'S OPPOSITION TO
MOTION FOR PRELIMINARY
INJUNCTION**

Date: November 13, 2023
Time: 1:30 p.m.
Ctrm: 5
Judge: Hon. William B.
Shubb
Trial Date: None set
Action Filed: 9/08/2023

23 I, Gabrielle D. Boutin, hereby declare as follows:

24 1. I am a Deputy Attorney General with the California
25 Department of Justice and serve as counsel in this action for
26 Defendant Attorney General Rob Bonta, in his official capacity.
27 I make this declaration in support of Defendant's Opposition to
28

1 Motion for Preliminary Injunction. I have personal, first-hand
 2 knowledge of the matters set forth below and, if called as a
 3 witness, I could and would testify competently thereto.

4 2. Attached hereto as Exhibits 1-12 are true and correct
 5 copies of all official legislative committee reports for Assembly
 6 Bill 587, 2021-2022 Reg. Sess. (Cal. 2022) that are available on
 7 the California Legislature's website. I downloaded the documents
 8 on October 25, 2023. They are also publicly available online at
 9 [https://leginfo.legislature.ca.gov/faces/billAnalysisClient.xhtml](https://leginfo.legislature.ca.gov/faces/billAnalysisClient.xhtml?bill_id=202120220AB587)
 10 [?bill_id=202120220AB587](https://leginfo.legislature.ca.gov/faces/billAnalysisClient.xhtml?bill_id=202120220AB587) (last accessed October 25, 2023).

11 The documents and exhibit numbers are as follows:

Exh.	Document
1.	Assem. Comm. on Privacy and Consumer Protection, Bill Analysis of Assembly Bill 587, as proposed to be amended, 2021-2022 Reg. Sess. (April 22, 2021)
2.	Assem. Judiciary Comm., Bill Analysis of Assembly Bill 587, as amended March 25, 2021, 2021-2022 Reg. Sess. (April 27, 2021)
3.	Assem. Appropriations Comm., Bill Analysis of Assembly Bill 587, as amended April 28, 2022, 2021-2022 Reg. Sess. (May 12, 2021)
4.	Assem. Floor Analysis of Assembly Bill 587, as amended April 28, 2021, 2021-2022 Reg. Sess. (May 24, 2021)
5.	Sen. Judiciary Comm., Bill Analysis of AB 587, as amended April 28, 2021, 2021-2022 Reg. Sess. (July 13, 2021)
6.	Sen. Judiciary Comm., Bill Analysis of AB 587, as amended June 23, 2022, 2021-2022 Reg. Sess. (June 28, 2022)
7.	Sen. Appropriations Comm., Bill Analysis of Assembly Bill 587, as amended June 30, 2022, 2021-2022 Reg. Sess. (Aug. 1, 2022)
8.	Sen. Appropriations Comm., Bill Analysis of Assembly Bill 587, Analysis Addendum-Suspense File, as amended Aug. 11, 2022, 2021-2022 Reg. Sess. (Aug. 11, 2022)
9.	Sen. Floor Analysis of Assembly Bill 587, as amended Aug. 11, 2022, 2021-2022 Reg. Sess. (Aug. 15, 2022)
10.	Sen. Floor Analysis of Assembly Bill 587, as amended Aug. 11, 2022, 2021-2022 Reg. Sess. (Aug. 26, 2022)

11.	Sen. Floor Analysis of Assembly Bill 587, as amended Aug. 24, 2022, 2021-2022 Reg. Sess. (Aug. 26, 2022)
12.	Assem. Floor Analysis of Assembly Bill 587, as amended Aug. 24, 2022, 2021-2022 Reg. Sess. (Aug. 30, 2022)

I declare under penalty of perjury under the laws of the State of California that the foregoing is true and correct and that this declaration was executed on October 27, 2023, in Davis, California.

/s/ Gabrielle D. Boutin
Gabrielle D. Boutin

Exhibit 1

Date of Hearing: April 22, 2021

ASSEMBLY COMMITTEE ON PRIVACY AND CONSUMER PROTECTION

Ed Chau, Chair

AB 587 (Gabriel) – As Amended March 25, 2021

AS PROPOSED TO BE AMENDED

SUBJECT: Social media companies: terms of service

SUMMARY: This bill would require social media companies, as defined, to post their terms of service in a manner reasonably designed to inform all users of specified policies and would require a social media company to submit quarterly reports, as specified, starting July 1, 2022, to the Attorney General (AG). Specifically, **this bill would:**

- 1) Require a social media company to post their terms of service (ToS) in a manner reasonably designed to inform all users of the internet-based service owned or operated by the social media company of the existence and contents of the terms of service, and require the ToS to include all of the following:
 - Contact information for the purpose of allowing users to ask the social media company questions about the terms of service.
 - A description of the process that users must follow to flag content, groups, or other users that they believe violate the terms of service, and the social media company's commitments on response and resolution time.
 - A list of potential actions the social media company may take against any item of content, or a user, or group of users, including, but not limited to, removal, demonetization, deprioritization, or banning.
- 2) Require the ToS to be available in all languages in which the social media company offers product features, including but not limited to menus and prompts.
- 3) Provide that a social media company is in violation of the above provisions if it fails to comply with the provisions of this section within 30 days of being notified of noncompliance by the AG.
- 4) Beginning July 1, 2022, require a social media company to submit to the AG, on a quarterly basis, a terms of service report, covering activity within the previous three months. The AG shall post on its website all ToS reports submitted pursuant to this requirement. The report shall include:
 - The current ToS of the social media company.
 - A complete and detailed description of any changes to the ToS since the last report, as specified;
 - A statement of whether the current version of the ToS defines each of the following categories of content, including their definitions, if applicable: (1) hate speech or racism;

- (2) extremism or radicalization; (3) disinformation or misinformation; (4) harassment; or (5) foreign political interference.
- A complete and detailed description of content moderation practices used by the social media company, including, but not limited to: (1) any existing policies intended to address the categories of content described immediately above; (2) any rules or guidelines regarding automated content moderation systems, as specified; (3) any training materials provided to content moderators, including educational materials; (4) responses to user reports of violations of the ToS; (5) any rules, guidelines, product changes and content moderator training materials that cover how the social media company would remove individual pieces of content, users, or groups that violate the ToS, or take broader action against individual users or against groups of users that violate the ToS; and (6) the languages in which the social media company offers product features, as specified.
 - Information, deidentified and disaggregated, as specified, on content that was flagged by the social media company as content belonging to any of the categories described above, including all of the following: (1) the total number of flagged items of content; (2) the total number of actioned items of content, including the total number of actioned items of content that were removed, demonetized, or deprioritized; (3) the number of times actioned items of content were viewed by users, shared, and the number of users that viewed that content before it was actioned; (4) the number of times users appealed social media company actions and reversals of those actions on appeal, as specified.
- 5) Provide that any violation of the above provisions shall be actionable under the Unfair Competition Law in addition to any other applicable state or federal law.
- 6) Provide various definitions including:
- “Actioned” to mean a social media company, due to a suspected or confirmed violation of the terms of service, that has taken some form of disciplinary action, including, but not limited to, removal, demonetization, deprioritization, or banning, against the relevant user or relevant item of content.
 - “Content” to mean media, including, but not limited to, text, images, videos, and groups of users that are created, posted, shared, or otherwise interacted with by users on an internet-based service.
 - “Social media company” to mean a person or entity that owns or operates a public-facing internet-based service that generated at least \$100,000,000 in gross revenue during the preceding year, and that allows users in the State to do all of the following: (1) construct a public or semipublic profile within a bounded system created by the service; (2) populate a list of other users with whom an individual shares a connection within the system; and (3) view and navigate a list of the individual’s connections and the connections made by other individuals within the system.
- 7) State that it is the intent of the Legislature that a social media company that violates this chapter shall be subject to meaningful remedies sufficient to induce compliance with the provisions above.

EXISTING LAW:

- 1) Provides, under the U.S. Constitution, that “Congress shall make no law . . . abridging the freedom of speech, or of the press, or the right of the people peaceably to assemble, and to petition the government for a redress of grievances.” (U.S. Const., 1st Amend., as applied to the states through the 14th Amendment’s Due Process Clause; *see Gitlow v. New York* (1925) 268 U.S. 652.)
- 2) Pursuant to the Communications Decency Act of 1996, provides, that “no provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider,” and affords broad protection from civil liability for the good faith content moderation decisions of interactive computer services. (47 U.S.C. Sec. 230(c)(1) and (2).)
- 3) Provides under the California Constitution for the right of every person to freely speak, write and publish his or her sentiments on all subjects, being responsible for the abuse of this right. Existing law further provides that a law may not restrain or abridge liberty of speech or press. (Cal. Const., art. I, Sec. 2(a).)
- 4) Establishes the Unfair Competition Law, which, among other things, provides for specific or preventive relief to enforce a penalty, forfeiture, or penal law in the case of unfair competition; and defines unfair competition to mean any unlawful, unfair or fraudulent business act or practice and unfair, deceptive, and untrue or misleading advertising. (Bus. & Prof. Code Sec. 17200, et seq.)
- 5) Permits actions for relief pursuant to 4), above, to be prosecuted exclusively by the Attorney General, a district attorney, a county counsel as specified, a city attorney as specified, or a city prosecutor as specified, in the name of the people of the State of California, or by a person who has suffered injury in fact and has lost money or property as a result of the unfair competition. (Bus. & Prof. Code Sec. 17204.)
- 6) Permits any person specified in 5), above, to seek injunctive relief and actual damages, and permits any person specified in 5) except for a person who has suffered injury in fact to pursue civil penalties, as specified, for violations of the provisions of the Unfair Competition Law. (Bus. & Prof. Code Secs. 17204 and 17206.)
- 7) Defines “social media” for the above proposes to mean an electronic service or account, or electronic content, including, but not limited to, videos, still photographs, blogs, video blogs, podcasts, instant and text messages, email, online services or accounts, or internet website profiles or locations. (Lab. Code Sec. 980(a).)

FISCAL EFFECT: Unknown

COMMENTS:

- 1) **Purpose of this bill:** This bill seeks to increase transparency and accountability with respect to content moderation policies on social media platforms by requiring social media companies to maintain ToS containing specified information, and mandating the submission of quarterly reports to the AG detailing content moderation policies and data related to content moderation practices and objectionable content. This bill is author sponsored.

2) **Author's statement:** According to the author:

In recent years, there has been growing concern around the role of social media in promoting hate speech, disinformation, conspiracy theories, violent extremism, and severe political polarization. Twitter, along with other social media platforms, has been implicated as a venue for hate groups to safely grow. A recent study of Twitter posts from 100 U.S. cities found that the greater proportion of tweets related to race- and ethnicity-based discrimination in a given city, the more hate crimes were occurring in that city. Robert Bowers, accused of murdering 11 elderly worshippers at a Pennsylvania synagogue in 2018, had been active on Gab, a Twitter-like site used by white supremacists. Most recently, investigations have shown that the violent riots at the Capitol in early January of this year were abetted and encouraged by posts on social media sites.

AB 587 would require social media platforms to publicly disclose their corporate policies and report key data and metrics around the enforcement of their policies. This disclosure would be accomplished through biannual and quarterly public filings with the Attorney General.

3) **Social media and content moderation:** As online social media become increasingly central to the public discourse, the companies responsible for managing social media platforms are faced with a complex dilemma regarding content moderation, i.e., how the platforms determine what content warrants disciplinary action such as removal of the item or banning of the user. In broad terms, there is a general public consensus that certain types of content, such as child pornography, depictions of graphic violence, emotional abuse, and threats of physical harm, are undesirable, and should be mitigated on these platforms to the extent possible. Many other categories of information, however, such as hate speech, racism, extremism, misinformation, political interference, and harassment, are far more difficult to reliably define, and assignment of their boundaries is often fraught with political bias. In such cases, both action and inaction by these companies seems to be equally maligned: too much moderation and accusations of censorship and suppressed speech arise; too little, and the platform risks fostering a toxic, sometimes dangerous community.

This dilemma has been at the forefront of the public conscience since, in the wake of the attack on the nation's capital on January 6, 2021, the sitting President of the United States was banned from some social media platforms for incitement of violence and propagation of misinformation. But the largest social media platforms are faced with thousands, if not millions of similarly difficult decisions related to content moderation on a daily basis. Despite the problem being more visible than ever, the machinations of content moderation in many ways remain a mystery. As a coalition of civil, minority, and immigrant rights organizations in support of the bill argues:

Despite the widespread nature of these concerns, efforts by social media companies to self-police such content have been opaque, arbitrary, biased, and inadequate. While some platforms share limited information about their efforts, the current lack of transparency has exacerbated concerns about the intent, enforcement, and impact of corporate policies, and deprived policymakers and the general public of critical data and metrics regarding the scope and scale of online hate and disinformation. Additional transparency is needed to allow consumers to make informed choices about the impact of these products

(including on their children) and so that researchers, civil society leaders, and policymakers can determine the best means to address this growing threat to our democracy.

AB 587 would address this troubling lack of transparency by requiring social media platforms to publicly disclose their corporate policies and report key data and metrics around the enforcement of their policies.

Efforts to address online content moderation at the state level have often been frustrated by issues of federal preemption. Specifically, Section 230 of the federal Communications Decency Act of 1996, which provides that an online platform generally cannot be held liable for content posted by third parties, explicitly preempts any conflicting state law. The law was designed to permit online platforms to freely moderate content in good faith without the risk of liability for content moderation decisions. But in effect, the liability shield provided by Section 230, coupled with its preemption of state law, makes it remarkably difficult to legislate at the state level with respect to content moderation. As a result, attempts to impose specific guidelines, restrictions, or requirements on social media platforms have thus far been unsuccessful.

AB 587 seeks to confront issues around social media content moderation practices by requiring the publication of ToS with specified information, and by requiring social media companies to submit quarterly reports containing information related to content moderation policies and data related to the application of those policies in practice.

- 4) **AB 587 would require social media companies to submit detailed reports on content moderation policies and practices to the AG:** AB 587 seeks to increase transparency with respect to content moderation policies and practices by large social media companies. Specifically, the bill consists of three main components: (1) ToS requirements; (2) reporting on content moderation policies and procedures; and (3) reporting on content moderation practices, including the data relating to the types of objectionable content being moderated.

The bill would require a social media company to post their ToS in a manner reasonably designed to inform users of their existence and contents, including in all languages in which the company offers product features, and would require those ToS to include all of the following information: (1) contact information for user inquiries relating to the ToS; (2) a description of the process users must follow to flag content, groups, or other users, that they believe violate the social media company's ToS, as well as commitments by the company with respect to response and resolution times for flagged items; and (3) a list of potential actions the social media company may take against an item of content, user, or group, including but not limited to removal, demonetization, deprioritization, or banning. The bill specifies that failure to comply with these provisions within 30 days of notification of noncompliance by the AG would constitute a violation.

Next, the bill would require a quarterly report to be submitted to the AG by the social media company consisting of specified information related to content moderation policies, including all of the following: (1) the current version of their ToS; (2) a complete and detailed description of any changes to the ToS since the previous report; (3) a statement of whether the ToS defines "hate speech or racism," "extremism or radicalization," "disinformation or misinformation," "harassment," or "foreign political interference," and if so, the definitions of those categories including any subcategories; (4) a complete and

detailed description of content moderation practices, including any policies intended to address categories described in (3), rules or guidelines regarding how automated content moderation systems enforce ToS and when and how those systems involve human review, training materials provided to content moderators, and a description of how the company responds to user reports of violations of the ToS; and (5) the languages in which the company offers product features and the languages for which the company has ToS.

Finally, the bill would require, in the same quarterly report, that the company provide information on content that was flagged by the company as content belonging to any of the categories described in (3), above, including all of the following: (1) the total number of flagged items of content; (2) the total number of actioned items of content, as defined; (3) the total number of actioned items of content that resulted in action taken by the company against the user or group of users responsible for the content; (4) the total number of actioned items of content that were removed, demonetized, or deprioritized by the social media company; (5) the number of times actioned items of content were viewed by users; (6) the number of times actioned items of content were shared, and the number of users that viewed the content before it was actioned; and (7) the number of times users appealed the company's actions and the number of reversals of the company's actions on appeal, disaggregated by each type of action. The bill would require that all such information be deidentified and disaggregated into the category of content, the type of content (e.g. posts, comments, messages, groups), the type of media of the content (e.g., text, images, videos), how the content was flagged (e.g., by human moderators, by AI software, by users), and how the content was actioned.

The bill would require the first of these reports to be submitted to the AG no later than July 1, 2022, and would require the AG to post on its official website all reports submitted pursuant to the bill.

Though content moderation on social media is a notoriously difficult problem to tackle, AB 587 seeming adopts a unique, data driven approach to progressing public policy in that space. Rather than placing specific content moderation requirements on companies, which in many cases raises constitutional issues, the bill instead provides for transparency and public accountability with respect to these practices, and establishes a timely, comprehensive dataset of untoward content on social media. This dataset can support research into the ever-changing social media ecosystem to help inform policies designed to root out its most problematic components while preserving its benefits for expression and connection.

- 5) **Opposition raises security and workability concerns regarding granularity of data required in reports:** AB 587 requires regular reporting by social media companies with respect to a wide range of information related to content moderation. Though granularity in this information can be useful for understanding the landscape and establishing transparency, opponents of the bill point out that too much granularity could put the platforms at risk. As a coalition of groups representing business interests argue in opposition to the bill:

In seeking to increase transparency around content moderation practices, AB 587 requires companies to report to the Attorney General the guidelines, practices, and even training materials companies use to moderate their platforms. This detailed information about content moderation practices, capabilities, and data regarding content moderation would not only threaten the security of these practices but provides bad actors with roadmaps to

get around our protections. We believe that while well intentioned, these requirements will ultimately allow scammers, spammers, and other bad actors to exploit our systems and moderators.

Indeed, in the past few years, the social media ecosystem has seen the emergence of sophisticated, sometimes state-sponsored actors seeking to exploit the design of their platforms toward nefarious ends. In this respect, it does not seem outlandish to presume that a large, detailed, public repository of information related to how content is moderated may increase sophistication of attempts to subvert content moderation systems. That said, in much the same way as policies for assessment and disclosure of security vulnerabilities are considered a best practice for cybersecurity, this same repository could enhance public scrutiny in a manner that would expose shortcomings in content moderation practices before they become catastrophic. Additionally, such information in aggregate from several platforms may facilitate comparison and meta-analysis that can help establish best practices that, even if transparent, are nonetheless secure. Accordingly, on balance, it is difficult to determine whether extensive, detailed publication of moderation practices would increase or decrease the vulnerability of these platforms to exploitation by bad actors.

Opponents of the bill further contend that granularity in the reporting of practical moderation data would be unworkable due to the magnitude of information that must be evaluated. The opposing coalition argues:

AB 587 requires businesses to report detailed metrics on a quarterly basis regarding not only the numerical scale of content moderation practices, but also details about how content is flagged and acted against. It would be nearly impossible to report this information quarterly due to the need to review, analyze, and adjudicate actioned content. Further, the sheer volume of content our companies review makes it similarly difficult to report data on individual pieces of content. Our companies address hundreds of millions of pieces of content across their platforms every few months. A requirement to collect, retain, and report information on individual pieces of content is unreasonable and unworkable. Furthermore, this volume of information is actually counterproductive to increasing transparency.

Notably, the provisions of this bill are limited to social media companies with over \$100,000,000 in gross revenue from the preceding year, a limitation intended to ensure that companies subject to the bill have ample resources to absorb the reporting requirement. Though the amount of content many of these platforms receive is indeed enormous, most platforms of this size and maturity internally perform detailed evaluation of content for product optimization purposes, necessitating the expertise and technical capacity to manage large datasets. In that sense, requiring management of such data to comply with reporting requirements would arguably be unlikely to exceed the capabilities of these companies. That said, considering the amount of individual items of content a single report would require if each item needed to be discussed individually, the reports, which are released for public consumption, would be virtually unreadable, and provide little benefit to the general public. While the previous version of the bill implied each item was to be dealt with independently, the author prudently amended the bill to deal with these items of actioned content collectively, only disaggregated across specified criteria.

- 6) **Liability and enforcement:** AB 587 specifies that a violation of its provisions is actionable under the Unfair Competition Law (UCL; Bus. & Prof. Code Sec. 17200) in addition to any other applicable state or federal law. The UCL creates a private right of action, but allows individual plaintiffs to seek only injunctive relief or restitutionary disgorgement, and only in the event the plaintiff can demonstrate injury-in-fact resulting from the violation. The UCL also permits the AG and district attorneys to bring causes of action in the name of the people of the State of California, and, in these cases, adds civil penalties up to \$2,500 per violation as an available remedy. Opponents of the bill express concerns that the liability exposure as a result of this enforcement mechanism may be counterproductive, and potentially unlawful. The coalition of business groups in opposition contends:

AB 587 opens companies up to the threat of liability and government investigation for routine moderation practices. Companies should not be subject to civil penalties or injunctive relief for the filing of a report, especially as comprehensive as the ones contemplated by this bill. Such litigation will deter investment in content moderation and suppress ongoing efforts to protect users from harmful content online. This extension of liability could also be interpreted to allow for lawsuits to be filed against platforms for the sufficiency of their moderation practices, which may be preempted by Section 230 of the Communications Decency Act (Section 230).

Staff notes that the bill does not appear to require any particular actions on the part of the company other than: (1) posting terms of service in accordance with specified criteria; and (2) submitting quarterly reports containing specified information. As such, it would appear that violations of the bill would only occur if the company failed to perform one or both of these requirements, and that so long as the reports and ToS conform to the specifications, the actual content moderation itself is not subject to enforcement. It therefore does not appear likely that liability imposed by this bill would allow for lawsuits to be filed against platforms for the sufficiency of their moderation practices, arguably making the risk of preemption on these grounds minimal.

That said, it is not clear whether the UCL is the appropriate mechanism for enforcing this bill, because it would be extremely difficult, if not impossible, for an individual to demonstrate injury-in-fact and loss of money or property as a result of a social media company's failure to submit a report or publish ToS. This leaves only public actions for injunctive relief or civil penalties. The bill does not make clear whether failure to submit a report in compliance with all specified requirements constitutes a single violation, or whether each non-compliant component is a separate violation. Assuming the former, the civil penalties available under the UCL are likely insufficient to enforce the bill, as complete noncompliance would result in a maximum annual penalty of \$12,500. For a company with gross annual revenue of over \$100,000,000, the threat of that penalty is not likely to ensure compliance. If the bill passes out of this Committee, the author may wish to consider amending the bill to provide for enforcement via specified civil penalties sufficient to ensure compliance.

- 7) **Related legislation:** AB 13 (Chau) would enact the Automated Decision Systems Accountability Act of 2021 and state the intent of the Legislature that state agencies use an acquisition method that minimizes the risk of adverse and discriminatory impacts resulting from the design and application of automated decision systems.

AB 35 (Chau) would social media platforms, as defined, to disclose whether or not that social media platform has a policy or mechanism in place to address the spread of misinformation, as specified. The bill would require the disclosure to be made easily accessible on the social media platform's website and mobile application.

AB 1379 (E. Garcia) would prohibit a social media platform from amplifying, in a manner that violates its terms of service or written public promises, content that is in violation of the platform's terms of service.

- 8) **Prior legislation:** AB 1316 (Gallagher, 2019) would have prohibited social media internet website operators located in California, as defined, from removing or manipulating content from that site on the basis of the political affiliation or political viewpoint of that content, except as specified. This bill was held in the Assembly Rules Committee.

AB 3169 (Gallagher, 2018) would have prohibited any person who operates a social media internet website or search engine located in California, as specified, from removing or manipulating content on the basis of the political affiliation or political viewpoint of that content. This bill failed passage in the Privacy & Consumer Protection Committee.

- 9) **Double referral:** This bill has been double-referred to the Assembly Judiciary Committee where it will be analyzed if passed by this Committee.

REGISTERED SUPPORT / OPPOSITION:

Support

AAUW Camarillo Branch
 American Jewish Committee - Los Angeles
 Anti-Defamation League
 Armenian Assembly of America
 Armenian National Committee of America - Western Region
 Buen Vecino
 California Asian Pacific American Bar Association
 California League of United Latin American Citizens
 Center for the Study of Hate & Extremism - California State University, San Bernardino
 Common Sense
 Hindu American Foundation, INC.
 Islamic Networks INC.
 Israeli-American Civic Action Network
 Japanese American Citizens League, Berkeley Chapter
 Jewish Center for Justice
 Jewish Public Affairs Committee
 Korean American Bar Association of Northern California
 Maplight
 National Hispanic Media Coalition
 Progressive Zionists of California
 Sikh American Legal Defense and Education Fund (SALDEF)
 Simon Wiesenthal Center, INC.
 Stonewall Democratic Club

Opposition

California Chamber of Commerce
Civil Justice Association of California
Internet Association
MPA - the Association of Magazine Media
TechNet

Analysis Prepared by: Landon Klein / P. & C.P. / (916) 319-2200

Exhibit 2

Date of Hearing: April 27, 2021

ASSEMBLY COMMITTEE ON JUDICIARY

Mark Stone, Chair

AB 587 (Gabriel) – As Amended March 25, 2021

As Proposed to be Amended

SUBJECT: SOCIAL MEDIA COMPANIES: TERMS OF SERVICE

KEY ISSUES:

- 1) SHOULD A SOCIAL MEDIA COMPANY BE REQUIRED TO POST TERMS OF SERVICE, ESPECIALLY AS TO THE ONLINE BEHAVIOR AND ACTIVITIES THAT IT EITHER PERMITS OR PROHIBITS?
- 2) SHOULD A SOCIAL MEDIA COMPANY BE REQUIRED TO SUBMIT QUARTERLY REPORTS DESCRIBING ITS CONTENT MODERATION POLICIES TO THE ATTORNEY GENERAL, AND SHOULD THE ATTORNEY GENERAL BE REQUIRED TO POST THOSE REPORTS ON ITS WEBSITE?

SYNOPSIS

This bill is one of three that the Committee had heard this year dealing with a social media company's ability to monitor and moderate posted content and users' online conduct. AB 35 (Chau) would require social media platforms to post what, if anything, they do to stop the spread of misinformation and violent or hateful content. AB 1114 (Gallagher) would require social media platforms to disclose how they address "unprotected" speech (obscenity, true threats, etc.); but for everything else, in its original form, AB 1114 would have made social media platforms a traditional public forum, thereby limiting their ability to remove objectionable content or deny access to certain users. The bill now before the Committee seeks transparency by requiring social media companies to post their "terms of service," especially as to the kinds of online behavior and activities they permit and, conversely, the kinds of behavior and activities that will lead the social media company to ban the user from its service. In addition, this bill would require social media companies to submit quarterly reports to the California Attorney General. In addition to including a copy of the company's terms of service, the reports would contain detailed descriptions of the methods used to moderate online activity. The report would also contain quantitative information, such as the number of times the social media company took action against users or content. Although the bill requires the Attorney General to post these reports on its website, it is not clear what else, if anything, the Attorney General would be expected to do with the information in the reports.

The bill is supported by several civil rights groups and ethnic and religious associations, who contend that social media has become a powerful and largely unregulated platform for groups espousing hate, violence, and bigotry. The bill is opposed by social media companies who fear that the information disclosed will allow "bad actors" to circumvent moderation control; they also believe the quarterly reports to the Attorney General are both burdensome and unnecessary. The bill passed out of the Assembly Privacy and Consumer Protection Committee on a 9-0 vote with an understanding that the author would take amendments in this Committee. Those

amendments are included in the summary below. However, the amendments have not removed the opposition.

SUMMARY: Requires a social media company to post terms of service (TOS) related to permitted and prohibited user behavior and activity and requires the social media company to submit reports and other relevant documents to the Attorney General (AG), as specified.

Specifically, **this bill:**

- 1) Requires a social media company to post their TOS in a manner reasonably designed to inform all users of the social media company's service of the existence and contents of the TOS, and require the TOS to include all of the following:
 - a) Contact information for the purpose of allowing users to ask the social media company questions about the terms of service.
 - b) A description of the process that users must follow to flag content, groups, or other users that they believe violate the terms of service, and the social media company's commitments on response and resolution time.
 - c) A list of potential actions the social media company may take against any item of content, or a user, or group of users, including, but not limited to, removal, demonetization, de-prioritization, or banning.
- 2) Requires the TOS to be available in all languages in which the social media company offers product features, including but not limited to menus and prompts.
- 3) Provides that a social media company is in violation of the above provisions if it fails to comply with the provisions of this section within 30 days of being notified of noncompliance by the AG.
- 4) Requires a social media company to submit to the AG, on a quarterly basis, a TOS report, covering activity within the previous three months. The AG shall post on its website all reports submitted pursuant to this requirement. The first report shall be submitted no later than July 1, 2022, and shall include:
 - a) A complete and detailed description of any changes to the TOS since the last report;
 - b) A statement of whether the current version of the TOS defines each of the following categories, including their definitions, if applicable: (1) hate speech or racism; (2) extremism or radicalization; (3) disinformation or misinformation; (4) harassment; or (5) foreign political interference;
 - c) A complete and detailed description of content moderation practices used by the social media company, including, but not limited to: (1) any existing policies intended to address the categories of content described immediately above; (2) any rules or guidelines regarding automated content moderation systems, as specified; (3) any training materials provided to content moderators, including educational materials; (4) responses to user reports of violations of the terms; (5) any rules, guidelines, product changes and content moderator training materials that cover how the social media company would remove individual pieces of content, users, or groups that violate the terms, or take

- broader action against individual users or against groups of users that violate the terms; and (6) the languages in which the social media company offers product features, as specified;
- d) Information, de-identified and disaggregated, as specified, on content that was flagged by the social media company as content belonging to any of the categories described above, including all of the following: (1) the total number of flagged items of content; (2) the total number of actioned items of content, including the total number of actioned items of content that were removed, demonetized, or deprioritized; (3) the number of times actioned items of content were viewed by users, shared, and the number of users that viewed that content before it was actioned; (4) the number of times users appealed social media company actions and reversals of those actions on appeal, as specified.
- 5) Provides that any violation of the above provisions shall be actionable under the Unfair Competition Law in addition to any other applicable state or federal law.
- 6) Provides definitions including:
- a) “Actioned” means a social media company, due to a suspected or confirmed violation of the TOS, has taken some form of disciplinary action, as specified.
- b) “Content” means media, including, but not limited to, text, images, videos, and groups of users that are created, posted, shared, or otherwise interacted with by users on an internet-based service.
- c) “Social media company” means a person or entity that owns or operates a public-facing internet-based service that generated at least \$100,000,000 in gross revenue during the preceding year, and that allows users in the State to do all of the following: (1) construct a public or semipublic profile within a bounded system created by the service; (2) populate a list of other users with whom an individual shares a connection within the system; and (3) view and navigate a list of the individual’s connections and the connections made by other individuals within the system.
- 7) States that it is the intent of the Legislature that a social media company that violates this chapter shall be subject to meaningful remedies sufficient to induce compliance with the provisions above.

EXISTING LAW:

- 1) Provides, under the Communications Decency Act of 1996, that “no provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider,” and affords broad protection from civil liability for the good faith content moderation decisions of interactive computer services. (47 U.S.C. Section 230(c)(1) and (2).)
- 2) Provides, under the Unfair Competition Law, for specific or preventive relief to enforce a penalty, forfeiture, or penal law in the case of unfair competition; and defines unfair competition to mean any unlawful, unfair or fraudulent business act or practice and unfair, deceptive, and untrue or misleading advertising. (Business & Professions Code Section 17200 *et seq.*)

- 3) Requires an operator of a commercial website or online service that collects personally identifiable information, as specified, to make its privacy policy available to consumers, as specified. (Civil Code Section 1798.130.)

FISCAL EFFECT: As currently in print this bill is keyed fiscal.

COMMENTS: According to the author:

Californians are becoming increasingly alarmed about the role of social media in promoting hate, disinformation, conspiracy theories, and extreme political polarization. Despite the widespread nature of these concerns, efforts by social media companies to self-police such content have been widely criticized as opaque, arbitrary, biased, and inadequate. While some platforms share limited information about their efforts, the current lack of transparency has exacerbated concerns about the intent, enforcement, and impact of corporate policies, and deprived policymakers and the general public of critical data and metrics regarding the scope and scale of online hate and disinformation. The public and policymakers deserve to know when and how social media companies are amplifying certain voices and silencing others. This is an important first step in a broader effort to protect our democracy and better regulate social media platforms.

What is the problem that this bill seeks to solve? According to the author and supporters of the bill, social media has become a powerful and largely unregulated platform for groups espousing hate, violence, bigotry, conspiracy theories, and disinformation. For example, the author cites a recent study of Twitter posts from 100 U.S. cities. The study found that the greater proportion of tweets related to race- and ethnicity-based discrimination in a given city, the more hate crimes were occurring in that city. The author also cites the case of Robert Bowers, the man who murdered eleven elderly worshipers at a Pennsylvania synagogue in 2018. Bowers had been active on Gab, a Twitter-like site used by white supremacists. The author contends that other investigations have suggested that the riots at the Capitol in early January of this year were encouraged by posts on social media sites.

Although this measure does not require social media companies to moderate or remove hateful or incendiary content, the author nonetheless sees the bill as “an important first step” in protecting our democracy from the dangerously divisive content that has become all too common on social media. In essence, AB 587 is a transparency measure. It would require social media companies to post their “terms of service,” especially as to the kinds of online behavior and activities they permit and, conversely, what kinds of behavior and activities subject a user to a temporary or permanent prohibition from using the social media service. Presumably, requiring these postings will serve three purposes. First, it will let users know what social media platforms do to flag and remove certain kinds of content, which may affect what sites users prefer to use. Second, it lets users know in advance what kind of content or conduct could lead to their being temporarily or permanently banned from using the social media service. Third, if social media companies are forced to disclose what they do in this regard, it may pressure them to become better corporate citizens by doing more to eliminate hate speech and disinformation.

Quarterly reports to the Attorney General. In addition to requiring social media companies to post terms of service for the user’s edification, this bill would also require social media companies to submit quarterly reports to the AG. In addition to including a copy of the company’s terms of services, the reports would also contain detailed descriptions of the methods that a company uses to moderate online activity. Specifically, the reports must contain a

“complete and detailed description” of the company’s policies, if any, for addressing hate speech or racism, extremism or radicalization, disinformation or misinformation, harassment, and foreign political interference. The report would also include detailed descriptions on a variety of other topics: the use of automated content moderation systems; the use of any training materials used to educate content moderators; how the company responds to user reports of violations of the terms of service; and any rules, guidelines, or training materials that cover how the company would remove individual pieces of content. The bill would also require the reports to contain quantitative information, including, among other things, the number of times that the social media company flagged or took action against content and the number of times that the company took action against users or groups of users responsible for the content. Finally, the reports would include information on *how* content was handled, including whether action was taken by employees, or by artificial intelligence software. The bill requires the AG to post these reports on its website. Other than positing these reports on its website, it is not clear what else, if anything, the Attorney General would be expected to do with the information in the reports. Although the bill does not yet prescribe a penalty for violations, it states the intent of the Legislature that violations of the bill’s requirements will be “subject to meaningful remedies sufficient to induce compliance,” presumably to be specified in the future.

Opposition concerns. Although some of the amendments taken by the author have assuaged some opposition concerns – such as eliminating a provision allowing the AG to demand documents from the social media companies – the social media companies, along with the California Chamber of Commerce, continue to oppose this measure.

Opposition concerns can be placed into two general categories. First, the opposition fears that the “complete and detailed” information required in the reports to the AG may include information that will allow “bad actors” to circumvent the social media companies content moderation policies – especially because the reports require detailed description on “training materials” and descriptions of the company’s use of automated systems and artificial intelligence. Second, the opposition questions the need for these reports. If the point of the bill is transparency – to inform consumers about the company’s content moderation rules and policies – then that objective is served by requiring the social media company to post its terms of service. Users, the opponents contend, are more likely to check the terms of service of a social media platform than they are to visit the AG’s website. Of course, the proponents of the bill could counter that the reports to the AG will contain *more* information than is required in the terms of service, including more detail on the methods used and numbers of times that the company has taken particular action. Opponents believe, however, that the cost of the reports is not justified by any additional information contained in the reports, which may not be any more useful to the user than the information already available in the company’s terms of service.

In light of the concerns raised by the opposition, the author may wish to consider, in particular, whether some of the information required in the reports – such as training manuals and descriptions of automated moderation systems – could, in fact, be exploited by those who want to circumvent a company’s moderation policies. In addition, the author may wish to consider the extent to which the information required in the terms of service mirrors the information required in the quarterly reports. If the information is not substantially different, then it is difficult to see what purpose is served by requiring the submission of the reports and their posting on a public website. Does the information provided in the AG reports provide information that is more useful to the user than that provided in the terms of service?

ARGUMENTS IN SUPPORT: This bill is supported by a wide coalition of groups who are particularly concerned about increased use of social media to promote hate, and sometimes incite violence, against groups based upon race, ethnicity, religion, or gender identity, among other factors. For example, the Anti-Defamation League (ADL), who writes on behalf of itself and several other organizations, supports AB 587 because it will “require social media platforms to publicly disclose their content moderation policies regarding online hate/racism, disinformation, extremism, and harassment, as well as regularly publish standardized key metrics and data around the enforcement of those policies.” In addition to citing studies suggesting an increase of online hate and extremism, ADL contends that “efforts by social media companies to self-police such content have been opaque, arbitrary, biased, and inadequate.” ADL believes that “AB 587 would address this troubling lack of transparency by requiring social media platforms to publicly disclose their corporate policies and report key data and metrics around the enforcement of their policies.”

Several other civil rights groups and ethnic and religious associations support this bill for substantially the same reasons as those articulated by the ADL.

ARGUMENTS IN OPPOSITION: Opponents of the measure, including the California Chambers of Commerce and associations representing internet-based and social media companies, claim that the disclosures required by this bill will expose social media moderation practices to “exploitation by bad actors.” In addition, opponents claim that the reporting requirements in the bill are “burdensome” and “unworkable.” While the opponents applaud the author’s “laudable goal of increasing transparency around content moderation practices,” they contend that the bill in print could ultimately “frustrate this goal.”

First, opponents contend that exposing too many details about content moderations practices could be exploited by “bad actors,” who would use that information to circumvent content moderation policies and raise new security threats. “To avoid undermining the goals of the bill and the work our companies have already undertaken to combat harmful content,” the opponents believe that the bill should not require the disclosure of such detailed information, especially the information on training materials, the automated content moderation systems used to enforce terms of service, and information on how the social media goes about removing content.

Second, in addition to potentially providing information to bad actors, opponents contend that the reporting requirements are “burdensome” and “will not advance the bill’s purpose.” Although the proposed amendments reduce the reporting requirements somewhat, the opponents contend that requiring “businesses to report detailed metrics on a quarterly basis” would be a nearly impossible task “due to the need to review, analyze, and adjudicate actioned content.” Moreover, “the sheer volume of content our companies review makes it similarly difficult to report data on an individual pieces of content.”

[Note: Other portions of the opposition letter concern provisions that have since been deleted from bill; nonetheless, opponents have informed that Committee that the proposed amendments do not eliminate their opposition, and, indeed, that adding a new provision requiring reports to contain a “complete and detailed description” of certain content moderation practices raises new concerns.]

REGISTERED SUPPORT / OPPOSITION:

Support

AAUW Camarillo Branch
American Jewish Committee - Los Angeles
Anti-defamation League
Armenian Assembly of America
Armenian National Committee of America - Western Region
California Asian Pacific American Bar Association
California League of United Latin American Citizens
Center for the Study of Hate & Extremism - California State University, San Bernardino
Common Sense
Hindu American Foundation, INC.
Islamic Networks INC.
Israeli-American Civic Action Network
Japanese American Citizens League, Berkeley Chapter
Jewish Center for Justice
Jewish Public Affairs Committee
Korean American Bar Association of Northern California
Maplight
National Hispanic Media Coalition
Progressive Zionists of California
Sikh American Legal Defense and Education Fund (SALDEF)
Simon Wiesenthal Center, INC.
Stonewall Democratic Club

Opposition

California Chamber of Commerce
Civil Justice Association of California
Internet Association
MPA the Association of Magazine Media
TechNet

Analysis Prepared by: Thomas Clark / JUD. / (916) 319-2334

Exhibit 3

Date of Hearing: May 12, 2021

ASSEMBLY COMMITTEE ON APPROPRIATIONS

Lorena Gonzalez, Chair

AB 587 (Gabriel) – As Amended April 28, 2021

Policy Committee:	Privacy and Consumer Protection	Vote:	9 - 0
	Judiciary		10 - 0

Urgency: No State Mandated Local Program: No Reimbursable: No

SUMMARY:

This bill requires a social media company to post its terms of service (TOS) related to permitted and prohibited user behavior and activity on its site. Specifically, this bill:

- 1) Requires a social media company to post its TOS in a manner reasonably designed to inform all users of the social media company's service of the existence and contents of the TOS and requires the TOS to include all of the following:
 - a) Contact information for the purpose of allowing users to ask the social media company questions about its TOS.
 - b) A description of the process users must follow to flag content, groups or other users they believe violate TOS, and the social media company's commitments on response and resolution time.
 - c) A list of potential actions the social media company may take against any item of content, or a user, or group of users, including, but not limited to, removal, demonetization and de-prioritization or banning.
- 2) Requires a social media company to submit to the Department of Justice (DOJ), on a quarterly basis, a TOS report, covering activity within the previous three months and requires the DOJ to post on its website all submitted reports with the first report submitted no later than July 1, 2022.
- 3) Provides that a social media company is in violation of this bill if it fails to comply with its requirements within 30 days of being notified of noncompliance by the AG.
- 4) States a violation of this bill is actionable under the Unfair Competition Law (UCL).

FISCAL EFFECT:

Costs (General Fund (GF)) possibly in the low to mid hundreds of thousands of dollars for the DOJ to review and post TOS reports on its website on a quarterly basis. Additional possibly significant cost pressures to the GF in the low millions of dollars in staff and resources, to the extent this bill results in the DOJ taking legal action against any social media company that does not comply. This bill states its intent to apply meaningful remedies sufficient to induce compliance, that a violation of the requirements of this bill are actionable under the UCL and that DOJ notify a social media company when it is not in compliance with the requirements of

this bill. DOJ currently enforces the UCL and other privacy laws and may be required to file for injunctive relief if a social media platform refuses to post its TOS as required by this bill.

COMMENTS:

1) **Purpose and Background.** According to the author:

The public and policymakers deserve to know when and how social media companies are amplifying certain voices and silencing others. This is an important first step in a broader effort to protect our democracy and better regulate social media platforms.

Section 230 of the Communications Decency Act of 1996 states online service providers or intermediaries that host or republish speech are protected against being legally responsible for what others say and do on online. Section 230 does not protect only internet service providers (ISPs), but also a range of “interactive computer service providers,” including basically any online service that publishes third-party content. Though there are important exceptions for certain criminal and intellectual property-based claims, Section 230 creates broad protections for social media platforms and may preempt any state law that would impose liability for content.

1) **Related Legislation.** AB 35 (Chau) requires a person that operates a social media platform to disclose what, if anything, it does to address the spread of misinformation. AB 35 is pending in this committee.

Analysis Prepared by: Kimberly Horiuchi / APPR. / (916) 319-2081

Exhibit 4

ASSEMBLY THIRD READING

AB 587 (Gabriel)

As Amended April 28, 2021

Majority vote

SUMMARY

This bill requires social media companies, as defined, to post their terms of service (ToS) in a manner reasonably designed to inform all users of specified policies and would require a social media company to submit quarterly reports, as specified, starting July 1, 2022, to the Attorney General (AG).

Major Provisions

- 1) Requires a social media company to post its ToS in a manner reasonably designed to inform all users of the social media company's service of the existence and contents of the ToS and requires the ToS to include all of the following:
 - a) Contact information for the purpose of allowing users to ask the social media company questions about its ToS.
 - b) A description of the process users must follow to flag content, groups or other users they believe violate ToS, and the social media company's commitments on response and resolution time.
 - c) A list of potential actions the social media company may take against any item of content, or a user, or group of users, including, but not limited to, removal, demonetization and de-prioritization or banning.
- 2) Requires a social media company to submit to the Department of Justice (DOJ), on a quarterly basis, a ToS report, covering content moderation policies and practices within the previous three months and requires the DOJ to post on its website all submitted reports with the first report submitted no later than July 1, 2022.
- 3) Provides that a social media company is in violation of this bill if it fails to comply with its requirements within 30 days of being notified of noncompliance by the AG.
- 4) States that a violation of this bill is actionable under the Unfair Competition Law (UCL).

COMMENTS

As online social media become increasingly central to the public discourse, the companies responsible for managing social media platforms are faced with a complex dilemma regarding content moderation, i.e., how the platforms determine what content warrants disciplinary action such as removal of the item or banning of the user. In broad terms, there is a general public consensus that certain types of content, such as child pornography, depictions of graphic violence, emotional abuse, and threats of physical harm, are undesirable, and should be mitigated on these platforms to the extent possible. Many other categories of information, however, such as hate speech, racism, extremism, misinformation, political interference, and harassment, are far more difficult to reliably define, and assignment of their boundaries is often fraught with political bias. In such cases, both action and inaction by these companies seems to be equally

maligned: too much moderation and accusations of censorship and suppressed speech arise; too little, and the platform risks fostering a toxic, sometimes dangerous community.

AB 587 seeks to confront issues around social media content moderation practices by requiring the publication of ToS with specified information, and by requiring social media companies to submit quarterly reports containing information related to content moderation policies and data related to the application of those policies in practice. Though content moderation on social media is a notoriously difficult problem to tackle, AB 587 seemingly adopts a unique, data driven approach to progressing public policy in that space. Rather than placing specific content moderation requirements on companies, which in many cases raises constitutional issues, the bill instead provides for transparency and public accountability with respect to these practices, and establishes a timely, comprehensive dataset of untoward content on social media. This dataset can support research into the ever-changing social media ecosystem to help inform policies designed to root out its most problematic components while preserving its benefits for expression and connection.

Though granularity in this information can be useful for understanding the landscape and establishing transparency, however, opponents of the bill point out that too much granularity could put the platforms at risk. Indeed, in the past few years, the social media ecosystem has seen the emergence of sophisticated, sometimes state-sponsored actors seeking to exploit the design of their platforms toward nefarious ends. In this respect, it does not seem outlandish to presume that a large, detailed, public repository of information related to how content is moderated may increase sophistication of attempts to subvert content moderation systems. That said, in much the same way as policies for assessment and disclosure of security vulnerabilities is considered a best practice for cybersecurity, this same repository could enhance public scrutiny in a manner that would expose shortcomings in content moderation practices before they become catastrophic. Additionally, such information in aggregate from several platforms may facilitate comparison and meta-analysis that can help establish best practices that, even if transparent, are nonetheless secure. Accordingly, on balance, it is difficult to determine whether extensive, detailed publication of moderation practices would increase or decrease the vulnerability of these platforms to exploitation by bad actors.

AB 587 specifies that a violation of its provisions is actionable under the Unfair Competition Law (UCL; Business & Professions Code Section 17200) in addition to any other applicable state or federal law. The UCL creates a private right of action, but allows individual plaintiffs to seek only injunctive relief or restitutionary disgorgement, and only in the event the plaintiff can demonstrate injury-in-fact resulting from the violation. The UCL also permits the AG and district attorneys to bring causes of action in the name of the people of the State of California, and, in these cases, adds civil penalties up to \$2,500 per violation as an available remedy. Opponents of the bill express concerns that the liability exposure as a result of this enforcement mechanism may be counterproductive, and potentially unlawful under Section 230 of the federal Communications Decency Act of 1996, which provides that an online platform generally cannot be held liable for content posted by third parties. Section 230 explicitly preempts any conflicting state law.

Staff notes that the bill does not appear to require any particular actions on the part of the company other than: 1) posting terms of service in accordance with specified criteria; and 2) submitting quarterly reports containing specified information. As such, it would appear that violations of the bill would only occur if the company failed to perform one or both of these

requirements, and that so long as the reports and ToS conform to the specifications, the actual content moderation itself is not subject to enforcement. It therefore does not appear likely that liability imposed by this bill would allow for lawsuits to be filed against platforms for the sufficiency of their moderation practices, arguably making the risk of preemption under Section 230 on these grounds minimal.

That said, it is not clear whether the UCL is the appropriate mechanism for enforcing this bill. Because it would be extremely difficult, if not impossible, for an individual to demonstrate injury-in-fact and loss of money or property as a result of a social media company's failure to submit a report or publish ToS, this leaves only public actions for injunctive relief or civil penalties. The bill does not make clear whether failure to submit a report in compliance with all specified requirements constitutes a single violation, or whether each non-compliant component is a separate violation. Assuming the former, the civil penalties available under the UCL are likely insufficient to enforce the bill, as complete noncompliance would result in a maximum annual penalty of \$12,500. For a company with gross annual revenue of over \$100,000,000, the threat of that penalty is not likely to ensure compliance.

According to the Author

In recent years, there has been growing concern around the role of social media in promoting hate speech, disinformation, conspiracy theories, violent extremism, and severe political polarization. Twitter, along with other social media platforms, has been implicated as a venue for hate groups to safely grow. A recent study of Twitter posts from 100 [United States] cities found that the greater proportion of tweets related to race- and ethnicity-based discrimination in a given city, the more hate crimes were occurring in that city. Robert Bowers, accused of murdering 11 elderly worshipers at a Pennsylvania synagogue in 2018, had been active on Gab, a Twitter-like site used by white supremacists. Most recently, investigations have shown that the violent riots at the Capitol in early January of this year were abetted and encouraged by posts on social media sites.

AB 587 would require social media platforms to publicly disclose their corporate policies and report key data and metrics around the enforcement of their policies. This disclosure would be accomplished through biannual and quarterly public filings with the Attorney General.

Arguments in Support

A coalition of civil, minority, and immigrant rights organizations including the Anti-Defamation League, Common Sense, and the California League of United Latin American Citizens argues:

[E]fforts by social media companies to self-police [problematic] content have been opaque, arbitrary, biased, and inadequate. While some platforms share limited information about their efforts, the current lack of transparency has exacerbated concerns about the intent, enforcement, and impact of corporate policies, and deprived policymakers and the general public of critical data and metrics regarding the scope and scale of online hate and disinformation. Additional transparency is needed to allow consumers to make informed choices about the impact of these products (including on their children) and so that researchers, civil society leaders, and policymakers can determine the best means to address this growing threat to our democracy.

AB 587 would address this troubling lack of transparency by requiring social media platforms to publicly disclose their corporate policies and report key data and metrics around the enforcement of their policies.

Arguments in Opposition

A coalition of groups representing business interests including CalChamber, Internet Association, and the Civil Justice Association of California argue in opposition unless amended:

In seeking to increase transparency around content moderation practices, AB 587 requires companies to report to the Attorney General the guidelines, practices, and even training materials companies use to moderate their platforms. This detailed information about content moderation practices, capabilities, and data regarding content moderation would not only threaten the security of these practices but provides bad actors with roadmaps to get around our protections. We believe that while well intentioned, these requirements will ultimately allow scammers, spammers, and other bad actors to exploit our systems and moderators. [...]

AB 587 opens companies up to the threat of liability and government investigation for routine moderation practices. Companies should not be subject to civil penalties or injunctive relief for the filing of a report, especially as comprehensive as the ones contemplated by this bill. Such litigation will deter investment in content moderation and suppress ongoing efforts to protect users from harmful content online. This extension of liability could also be interpreted to allow for lawsuits to be filed against platforms for the sufficiency of their moderation practices, which may be preempted by Section 230 of the Communications Decency Act (Section 230).

FISCAL COMMENTS

According to the Assembly Appropriations Committee, "[c]osts (General Fund (GF)) possibly in the low to mid hundreds of thousands of dollars for the DOJ to review and post TOS reports on its website on a quarterly basis. Additional possibly significant cost pressures to the GF in the low millions of dollars in staff and resources, to the extent this bill results in the DOJ taking legal action against any social media company that does not comply. This bill states its intent to apply meaningful remedies sufficient to induce compliance, that a violation of the requirements of this bill are actionable under the UCL and that DOJ notify a social media company when it is not in compliance with the requirements of this bill. DOJ currently enforces the UCL and other privacy laws and may be required to file for injunctive relief if a social media platform refuses to post its TOS as required by this bill."

VOTES**ASM PRIVACY AND CONSUMER PROTECTION: 9-0-2**

YES: Chau, Bauer-Kahan, Bennett, Carrillo, Cunningham, Gabriel, Irwin, Lee, Wicks

ABS, ABST OR NV: Kiley, Gallagher

ASM JUDICIARY: 10-0-1

YES: Stone, Gallagher, Chau, Chiu, Davies, Lorena Gonzalez, Holden, Kalra, Maienschein, Reyes

ABS, ABST OR NV: Kiley

ASM APPROPRIATIONS: 13-0-3

YES: Lorena Gonzalez, Calderon, Carrillo, Chau, Davies, Gabriel, Eduardo Garcia, Levine, Quirk, Robert Rivas, Akilah Weber, Holden, Luz Rivas

ABS, ABST OR NV: Bigelow, Megan Dahle, Fong

UPDATED

VERSION: April 28, 2021

CONSULTANT: Landon Klein / P. & C.P. / (916) 319-2200

FN: 0000489

Exhibit 5

SENATE JUDICIARY COMMITTEE
Senator Thomas Umberg, Chair
2021-2022 Regular Session

AB 587 (Gabriel)
Version: April 28, 2021
Hearing Date: July 13, 2021
Fiscal: Yes
Urgency: No
CK

SUBJECT

Social media companies: terms of service

DIGEST

This bill requires social media companies, as defined, to post their terms of service and to submit quarterly reports to the Attorney General on their terms of service and content moderation policies and outcomes.

EXECUTIVE SUMMARY

In 2005, five percent of adults in the United States used social media. In just six years, that number jumped to half of all Americans. Today, over 70 percent of adults use at least one social media platform. Facebook alone is used by 69 percent of adults, and 70 percent of those adults say they use the platform on a daily basis.

Given the reach of social media platforms and the role they play in many people's lives, concerns have arisen over what content permeates these sites, entering the lives of the billions of users, and the effects that has on them and society as a whole. In particular, the sharpest calls for action focus on the rampant spread of misinformation, hate speech, and sexually explicit content. Social media companies' content moderation of a decade ago involved handfuls of individuals and user policies were minimal. These programs and policies have dramatically evolved over the years but the proliferation of objectionable content and "fake news" has led to calls for swifter and more aggressive action in response. However, there has also been backlash against perceived censorship in response to filtering of content and alleged "shadow banning."

This bill requires social media companies, as defined, to publicly post their terms of service, with certain required elements, and to provide the Attorney General with a quarterly report on their content moderation procedures and outcomes.

This bill is sponsored by the Anti-Defamation League. It is supported by a variety of groups, including Common Sense and the Islamic Networks Group. It is opposed by various technology and business associations, including the California Chamber of Commerce, the Internet Association, and TechNet.

PROPOSED CHANGES TO THE LAW

Existing law:

- 1) Prohibits, through the United States Constitution, the enactment of any law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances. (U.S. Const. Amend. 1.)
- 2) Provides, through the California Constitution, for the right of every person to freely speak, write and publish their sentiments on all subjects, being responsible for the abuse of this right. Existing law further provides that a law may not restrain or abridge liberty of speech or press. (Cal. Const., art. I, § 2(a).)
- 3) Provides, in federal law, that a provider or user of an interactive computer service shall not be treated as the publisher or speaker of any information provided by another information content provider. (47 U.S.C. § 230(c)(2).)
- 4) Provides that a provider or user of an interactive computer service shall not be held liable on account of:
 - a) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected; or
 - b) any action taken to enable or make available to information content providers or others the technical means to restrict access to such material. (47 U.S.C. § 230(c)(2).)
- 5) Defines “interactive computer service” as any information service, system, or access software provider that provides or enables computer access by multiple users to a computer server, including specifically a service or system that provides access to the Internet and such systems operated or services offered by libraries or educational institutions. (47 U.S.C. § 230(f)(2).)
- 6) Establishes the Unfair Competition Law (UCL) and defines “unfair competition” to mean and include any unlawful, unfair, or fraudulent business act or practice and unfair, deceptive, untrue, or misleading advertising and any act prohibited

by Chapter 1 (commencing with Section 17500) of Part 3 of Division 7 of the Business and Professions Code. (Bus. & Prof. Code § 17200 et seq.)

- 7) Provides that any person who engages, has engaged, or proposes to engage in unfair competition may be enjoined. Any person may pursue representative claims or relief on behalf of others only if the claimant meets specified standing requirements and complies with Section 382 of the Code of Civil Procedure, but these limitations do not apply to claims brought under this chapter by the Attorney General, or any district attorney, county counsel, city attorney, or city prosecutor in this state. (Bus. & Prof. Code § 17203.)
- 8) Requires actions for relief pursuant to the UCL be prosecuted exclusively in a court of competent jurisdiction and only by the following:
 - a) the Attorney General;
 - b) a district attorney;
 - c) a county counsel authorized by agreement with the district attorney in actions involving violation of a county ordinance;
 - d) a city attorney of a city having a population in excess of 750,000;
 - e) a city attorney in a city and county;
 - f) a city prosecutor in a city having a full-time city prosecutor in the name of the people of the State of California upon their own complaint or upon the complaint of a board, officer, person, corporation, or association with the consent of the district attorney; or
 - g) a person who has suffered injury in fact and has lost money or property as a result of the unfair competition. (Bus. & Prof. Code § 17204.)
- 9) Holds any person who engages, has engaged, or proposes to engage in unfair competition liable for a civil penalty not to exceed \$2,500 for each violation, which shall be assessed and recovered in a civil action brought by the Attorney General, or other public prosecutors. (Bus. & Prof. Code § 17206(a).)
- 10) Prohibits false or deceptive advertising to consumers about the nature of any property, product, or service, including false or misleading statements made in print, over the internet, or any other advertising method. (Bus. & Prof. Code § 17500.)
- 11) Defines libel as a false and unprivileged publication by writing, printing, or any other representation that exposes any person to hatred, contempt, ridicule, or obloquy, which causes that person to be shunned or avoided, or which has a tendency to injure that person in their occupation. (Civ. Code §§ 45, 47.)
- 12) Requires certain businesses to disclose the existence and details of specified policies, including:

- a) Operators of commercial websites or online services that collect personally identifiable information about individual consumers residing in California who use or visit the website must conspicuously post its privacy policy. (Bus. & Prof. Code § 22575.)
- b) Retailers and manufacturers doing business in this state and having annual worldwide gross receipts over \$100,000,000 must disclose online whether the business has a policy to combat human trafficking and, if so, certain details about that policy. (Civ. Code § 1714.43.)
- c) End-users of automated license plate recognition technology must post its usage and privacy policy on its website. (Civ. Code § 1798.90.53.)
- d) Campus bookstores at public postsecondary educational institutions must post in-store or online a disclosure of its retail pricing policy on new and used textbooks. (Educ. Code § 66406.7(f).)

This bill:

- 1) Requires a social media company to post their terms of service in a manner reasonably designed to inform all users of the internet-based service owned or operated by the social media company of the existence and contents of the terms of service. The terms of service shall include all of the following:
 - a) contact information for the purpose of allowing users to ask the social media company questions about the terms of service;
 - b) a description of the process that users must follow to flag content, groups, or other users that they believe violate the terms of service, and the social media company's commitments on response and resolution time; and
 - c) a list of potential actions the social media company may take against an item of content or a user, including, but not limited to, removal, demonetization, deprioritization, or banning.
- 2) Requires the terms of service to be available in all languages in which the social media company offers product features, including, but not limited to, menus and prompts.
- 3) Provides that a social media company shall be in violation only if the social media company fails to comply within 30 days of being notified of noncompliance by the Attorney General.
- 4) Requires social media companies to submit a terms of service report, quarterly, with the first report due July 1, 2022, to the Attorney General, who must post it on their website. The terms of service report must include the following:
 - a) the current version of the terms of service of the social media company;
 - b) if a social media company has filed its first quarterly report, a complete and detailed description of any changes to the terms of service since the last quarterly report;

- c) a statement of whether the current version of the terms of service defines specified categories of content, and, if so, the definitions of those categories, including any subcategories. This includes hate speech, racism, extremism, harassment, disinformation, and foreign political interference;
 - d) a complete and detailed description of content moderation practices used by the social media company, including, but not limited to, all of the following:
 - i. any policies intended to address the above categories of content;
 - ii. any rules or guidelines regarding how a social media company's automated content moderation systems enforce terms of service and when these systems involve human review;
 - iii. any training materials provided to human content moderators intended to educate them on the above categories of content;
 - iv. how the social media company responds to user reports of violations of the terms of service;
 - v. any rules, guidelines, product changes, and content moderator training materials that cover how the social media company would remove individual pieces of content, users, or groups that violate the terms of service, or take broader action against individual users or against groups of users that violate the terms of service;
 - vi. the languages in which the social media company offers product features, and the languages for which the social media company has terms of service;
 - e) information on content that was flagged by the social media company as content belonging to any of the above categories, including the total number of all of the following:
 - i. flagged items of content;
 - ii. actioned items of content;
 - iii. actioned items of content that resulted in action taken by the social media company against the user or users responsible;
 - iv. actioned items of content that were removed, demonetized, or deprioritized by the social media company;
 - v. times actioned items of content were viewed by users;
 - vi. times actioned items of content were shared, and the number of users that viewed the content before it was actioned; and
 - vii. times users appealed social media company actions and the number of reversals on appeal disaggregated by each action;
 - f) all information required by (e) shall also be disaggregated into the category of content, the type of content, the type of media, and how the content was flagged and actioned.
- 5) Defines "social media company" as a person or entity that owns or operates a public-facing internet-based service that generated at least \$100,000,000 in gross

revenue during the preceding calendar year, and that allows users in the state to do all of the following:

- a) construct a public or semipublic profile within a bounded system created by the service;
 - b) populate a list of other users with whom an individual shares a connection within the system; and
 - c) view and navigate a list of the individual's connections and the connections made by other individuals within the system.
- 6) Provides that a "social media company" does not include a person or entity that exclusively owns and operates an electronic mail service.
- 7) Defines "actioned" to mean a social media company, due to a suspected or confirmed violation of the terms of service, has taken some form of action, including, but not limited to, removal, demonetization, deprioritization, or banning, against the relevant user or relevant item of content.
- 8) Defines "terms of service" as a policy adopted by a social media company that specifies, at least, the user behavior and activities that are permitted on the internet-based service owned or operated by the social media company, and the user behavior and activities that may subject the user or an item of content to being actioned. This may include, but is not limited to, a terms of service document or agreement, rules or content moderation guidelines, community guidelines, acceptable uses, and other policies and established practices that outline these policies.
- 9) Makes violations of its provisions actionable under the Unfair Competition Law, Business and Professions Code section 17200 et seq., and any other applicable state or federal law.

COMMENTS

1. Social media content

In recent years, the clamor for more robust content moderation on social media has reached a fever pitch. This includes calls to control disinformation or "fake news," hate speech, political interference, and other online harassment.

The 2016 election was a major breaking point for many. Investigations uncovered attempted interference in the United States Presidential election through a social media "information warfare campaign designed to spread disinformation and societal division

in the United States.”¹ The United States Senate Select Committee on Intelligence issued a report detailing how Russian operatives carried out their plan:

Masquerading as Americans, these operatives used targeted advertisements, intentionally falsified news articles, self-generated content, and social media platform tools to interact with and attempt to deceive tens of millions of social media users in the United States. This campaign sought to polarize Americans on the basis of societal, ideological, and racial differences, provoked real world events, and was part of a foreign government's covert support of Russia's favored candidate in the U.S. presidential election.

This again became a threat in the 2020 election, with social media rife with misinformation such as the incorrect election date,² and then social media became a hotbed of misinformation about the results of the election.³ The author points to investigations that have found the violent insurrectionists that stormed the Capitol on January 6, 2021, were abetted and encouraged by posts on social media sites.⁴ In response to indications that social media provided a venue for those who overran and assaulted police officers, Facebook deflected blame, asserting that “these events were largely organized on platforms that don’t have our abilities to stop hate, don’t have our standards, and don’t have our transparency.”⁵ However, later indictments of those perpetrating the attack “made it clear just how large a part Facebook had played, both in spreading misinformation about election fraud to fuel anger among the Jan. 6 protesters, and in aiding the extremist militia’s communication ahead of the riots.”⁶

¹ Select Committee on Intelligence, Russian Active Measures, Campaigns, and Interference in the 2016 U.S. Election, United States Senate, https://www.intelligence.senate.gov/sites/default/files/documents/Report_Volume2.pdf. All internet citations are current as of July 8, 2021.

² Pam Fessler, *Robocalls, Rumors And Emails: Last-Minute Election Disinformation Floods Voters*, NPR (October 24, 2020), <https://www.npr.org/2020/10/24/927300432/robocalls-rumors-and-emails-last-minute-election-disinformation-floods-voters>.

³ Sheera Frenkel, *How Misinformation ‘Superspreaders’ Seed False Election Theories*, New York Times (November 23, 2020), <https://www.nytimes.com/2020/11/23/technology/election-misinformation-facebook-twitter.html>; Philip Bump, *The chain between Trump’s misinformation and violent anger remains unbroken*, Washington Post (May 12, 2021), <https://www.washingtonpost.com/politics/2021/05/12/chain-between-trumps-misinformation-violent-anger-remains-unbroken/>.

⁴ Ken Dilanian & Ben Collins, *There are hundreds of posts about plans to attack the Capitol. Why hasn't this evidence been used in court?* (April 20, 2021) NBC News, <https://www.nbcnews.com/politics/justice-department/we-found-hundreds-posts-about-plans-attack-capitol-why-aren-n1264291>.

⁵ Sheera Frenkel & Cecilia Kang, *Mark Zuckerberg and Sheryl Sandberg’s Partnership Did Not Survive Trump* (July 8, 2021) The New York Times, <https://www.nytimes.com/2021/07/08/business/mark-zuckerberg-sheryl-sandberg-facebook.html>.

⁶ *Ibid.*

One area the author specifically focuses in on as motivation for the bill is the rise of hate speech online and the real world consequences. The author points to a recent study of over 500 million Twitter posts from 100 cities in the United States that found that “more targeted, discriminatory tweets posted in a city related to a higher number of hate crimes.”⁷

Misinformation also poses a danger to public health: One study found that the more people rely on social media as their main news source, the more likely they are to believe misinformation about the COVID-19 pandemic.⁸ Another found that a mere 12 people are responsible for 65 percent of the false and misleading claims about COVID-19 vaccines on Facebook, Instagram, and Twitter.⁹ Misinformation hinders emergency responses to natural responses, when social media posts contain incorrect or out-of-date information.¹⁰

The author frames the problem:

Over the past several years, there has been growing concern around the role of social media in promoting hate speech, disinformation, conspiracy theories, violent extremism, and severe political polarization. If properly managed, the ability for social media to amplify ideas and messages that would otherwise lack widespread exposure can give voice to otherwise marginalized populations and improve the public discourse, but the same capacity can feed the propagation of misinformation and dangerous rhetoric.

Writing in support, the Anti-Defamation League, the sponsor of this bill, further explains the context of the bill:

In recent years, there has been growing concern around the role of social media in promoting hate speech, disinformation, conspiracy theories, violent extremism, harassment, and severe political polarization.

⁷ Press Release, *Hate speech on Twitter predicts frequency of real-life hate crimes* (June 24, 2019) NYU Tandon School of Engineering, <https://engineering.nyu.edu/news/hate-speech-twitter-predicts-frequency-real-life-hate-crimes>.

⁸ Yan Su, *It doesn't take a village to fall for misinformation: Social media use, discussion heterogeneity preference, worry of the virus, faith in scientists, and COVID-19-related misinformation belief* (May 2021) *Telematics and Information*, Vol. 58, <https://www.sciencedirect.com/science/article/abs/pii/S0736585320302069?via%3Dihub>.

⁹ Shannon Bond, *Just 12 People Are Behind Most Vaccine Hoaxes On Social Media, Research Shows* (May 14, 2021) NPR, <https://www.npr.org/2021/05/13/996570855/disinformation-dozen-test-facebooks-tweeters-ability-to-curb-vaccine-hoaxes>.

¹⁰ United States Department of Homeland Security, *Countering False Information on Social Media in Disasters and Emergencies* (March 2018), https://www.dhs.gov/sites/default/files/publications/SMWG_Countering-False-Info-Social-Media-Disasters-Emergencies_Mar2018-508.pdf.

According to ADL's 2021 Online Hate and Harassment Survey, 41% of individuals experience online harassment and one in three of those individuals attribute at least some harassment to their identity. Identity-based harassment remains worrisome, affecting the ability of already marginalized communities to be safe in digital spaces.

Importantly, this hate and harassment isn't only taking place in the dark corners of the internet. 75% of ADL's 2021 Online Hate and Harassment Survey respondents who were harassed said at least some harassment happened on Facebook – and many also attributed harassment to other mainstream social media platforms. And online extremism is also front and center: Facebook's own researchers found that 64% of people who joined an extremist group on Facebook only did so because the company's algorithm recommended it to them.

A recent Congressional Research Services Report discussed the issue of content moderation and specifically the spread of misinformation and the role that social media companies play in worsening the issue:

Two features of social media platforms—the user networks and the algorithmic filtering used to manage content—can contribute to the spread of misinformation. Users can build their own social networks, which affect the content that they see, including the types of misinformation they may be exposed to. Most social media operators use algorithms to sort and prioritize the content placed on their sites. These algorithms are generally built to increase user engagement, such as clicking links or commenting on posts. In particular, social media operators that rely on advertising placed next to user-generated content as their primary source of revenue have incentives to increase user engagement. These operators may be able to increase their revenue by serving more ads to users and potentially charging higher fees to advertisers. Thus, algorithms may amplify certain content, which can include misinformation, if it captures users' attention.¹¹

The role that content moderation, or the lack of it, has in alleviating or exacerbating these issues has been a source of much debate. A policy paper published by the Shorenstein Center on Media, Politics, and Public Policy at the Harvard Kennedy School, *Countering Negative Externalities in Digital Platforms*, focuses on the costs associated with various internet platforms that are not absorbed by the companies themselves:

¹¹ Jason A. Gallo & Clare Y. Cho, *Social Media: Misinformation and Content Moderation Issues for Congress* (January 27, 2021) Congressional Research Service, <https://crsreports.congress.gov/product/pdf/R/R46662>.

Today, in addition to the carcinogenic effects of chemical runoffs and first and second hand tobacco smoke, we have to contend with a new problem: the poisoning of our democratic system through foreign influence campaigns, intentional dissemination of misinformation, and incitements to violence inadvertently enabled by Facebook, YouTube and our other major digital platform companies.¹²

The paper asserts that these major platform companies “enable exceptionally malign activities” and “experience shows that the companies have not made sufficient investments to eliminate or reduce these negative externalities.”

As pointed out by recent Wall Street Journal reporting, the companies’ employees are aware of the dangers:

A Facebook Inc. team had a blunt message for senior executives. The company’s algorithms weren’t bringing people together. They were driving people apart.

“Our algorithms exploit the human brain’s attraction to divisiveness,” read a slide from a 2018 presentation. “If left unchecked,” it warned, Facebook would feed users “more and more divisive content in an effort to gain user attention & increase time on the platform.”

That presentation went to the heart of a question dogging Facebook almost since its founding: Does its platform aggravate polarization and tribal behavior?

The answer it found, in some cases, was yes.¹³

A recent New York Times article on leadership at Facebook elaborates:

To achieve its record-setting growth, the [Facebook] had continued building on its core technology, making business decisions based on how many hours of the day people spent on Facebook and how many times a day they returned. Facebook’s algorithms didn’t measure if the magnetic force pulling them back to Facebook was the habit of wishing a friend happy birthday, or a rabbit hole of conspiracies and misinformation.

¹² *Countering Negative Externalities in Digital Platforms* (October 7, 2019) Shorenstein Center on Media, Politics and Public Policy, <https://shorensteincenter.org/countering-negative-externalities-in-digital-platforms/>.

¹³ Jeff Horowitz & Deepa Seetharaman, *Facebook Executives Shut Down Efforts to Make the Site Less Divisive* (May 26, 2020) Wall Street Journal, <https://www.wsj.com/articles/facebook-knows-it-encourages-division-topexecutives-nixed-solutions-11590507499>.

Facebook's problems were features, not bugs.¹⁴

Another paper recently released provides "Recommendations to the Biden Administration," and is relevant to the considerations here:

The Administration should work with Congress to develop a system of financial incentives to encourage greater industry attention to the social costs, or "externalities," imposed by social media platforms. A system of meaningful fines for violating industry standards of conduct regarding harmful content on the internet is one example. In addition, the Administration should promote greater transparency of the placement of digital advertising, the dominant source of social media revenue. This would create an incentive for social media companies to modify their algorithms and practices related to harmful content, which their advertisers generally seek to avoid.¹⁵

2. Content moderation, transparency, and the low-grade war on our cognitive security

There are a number of considerations when addressing how to approach the proliferation of these undesirable social media posts and the companies' practices that fuel the flames. A number of methods of content moderation are being deployed and have evolved from simply blocking content or banning accounts to quarantining topics, removing posts from search results, barring recommendations, and down ranking posts in priority. However, there is a lack of transparency and understanding of exactly what companies are doing and why it does not seem to be enough. A recent article in the MIT Technology Review articulates the issues with content moderation behind the curtain:

As social media companies suspended accounts and labeled and deleted posts, many researchers, civil society organizations, and journalists scrambled to understand their decisions. The lack of transparency about those decisions and processes means that—for many—the election results end up with an asterisk this year, just as they did in 2016.

What actions did these companies take? How do their moderation teams work? What is the process for making decisions? Over the last few years,

¹⁴ Sheera Frenkel & Cecilia Kang, *Mark Zuckerberg and Sheryl Sandberg's Partnership Did Not Survive Trump* (July 8, 2021) The New York Times, <https://www.nytimes.com/2021/07/08/business/mark-zuckerberg-sheryl-sandberg-facebook.html>.

¹⁵ Caroline Atkinson, et al., *Recommendations to the Biden Administration On Regulating Disinformation and Other Harmful Content on Social Media* (March 2021) Harvard Kennedy School & New York University Stern School of Business, https://static1.squarespace.com/static/5b6df958f8370af3217d4178/t/6058a456ca24454a73370dc8/1616421974691/TechnologyRecommendations_2021final.pdf.

platform companies put together large task forces dedicated to removing election misinformation and labeling early declarations of victory. Sarah Roberts, a professor at UCLA, has written about the invisible labor of platform content moderators as a shadow industry, a labyrinth of contractors and complex rules which the public knows little about. Why don't we know more?

In the post-election fog, social media has become the terrain for a low-grade war on our cognitive security, with misinformation campaigns and conspiracy theories proliferating. When the broadcast news business served the role of information gatekeeper, it was saddled with public interest obligations such as sharing timely, local, and relevant information. Social media companies have inherited a similar position in society, but they have not taken on those same responsibilities. This situation has loaded the cannons for claims of bias and censorship in how they moderated election-related content.

This bill seeks to increase transparency around what terms of service social media companies are setting out and how it ensures those terms are abided by. The goal is to learn more about the methods of content moderation and how successful they are. According to the author:

The line between providing an open forum for productive discourse and permitting the proliferation of hate speech and misinformation is a fine one, and depends largely on the structure and practices of the platform. However, these platforms rarely provide detailed insight into such practices, and into the relative effectiveness of different approaches. This, along with constraints imposed by existing federal law, has historically made policy-making in this space remarkably difficult. This bill seeks to provide critical transparency to both inform the public as to the policies and practices governing the content they post and engage with on social media, and to allow for comparative assessment of content moderation approaches to better equip both social media companies and policymakers to address these growing concerns.

ADL emphasizes the need for the bill:

Despite the widespread nature of these concerns, efforts by social media companies to self-police such content have been opaque, arbitrary, biased, and inadequate. While some platforms share limited information about their efforts, the current lack of transparency has exacerbated concerns about the intent, enforcement, and impact of corporate policies. Consequently, policymakers and the general public remain deprived of critical data and metrics regarding the scope and scale of online hate and

disinformation. Additional transparency is needed to allow consumers to make informed choices about the impact of these products (including on their children) and so that researchers, civil society leaders, and policymakers can take meaningful action to decrease online hate and extremism, and to address this growing threat to our democracy.

The creation of a thoughtful and standardized enumeration and measurement of policies and enforcement will serve policymakers and the public. We need this critical information to better understand the policies and practices of social media platforms – which have a profound impact on communication and discourse.

AB 587 will address this troubling lack of transparency by requiring social media platforms to publicly disclose their corporate policies and report key data and metrics around the enforcement of their policies. This disclosure would be accomplished through regular public filings with the Attorney General.

This bill starts with a baseline requirement to have social media companies post their terms of service. These policies must include information about how users can ask questions, how they can flag content or users in violation, and a list of potential actions that the company might take in response. To ensure meaningful access, the terms of service must be posted in a manner reasonably designed to inform all users of their existence and contents and available in all languages in which the company offers product features. The Attorney General must provide a 30-day right to cure before taking action against companies for failing to abide by these requirements.

The bill next requires an extremely detailed report to be compiled by these companies and submitted to the Attorney General on a quarterly basis. This report must include information on the terms of service, any changes made and whether they define certain categories of content, including hate speech or racism; extremism or radicalization; disinformation or misinformation; harassment; and foreign political interference.

The bill also requires the report to contain a “complete and detailed description of content moderation practices” used by the company. There must also be outcome-focused information included. Companies must report on the number of flagged items of content and the number of times the company took action in response. To understand the impact of the reported content, the report must detail the number of times this content was viewed and shared by users. The data must also include these details broken down by content category, the type of media, and other factors.

As the author references above, all of this occurs within tight quarters due to federal statutory and constitutional law. Section 230 of the Communications Decency Act, in relevant part, immunizes providers from liability for actions taken in good faith “to

restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected.” (47 U.S.C. § 230.) It also makes clear that platforms cannot “be treated as the publisher or speaker of any information provided by another information content provider.” Section 230’s language is intended to provide a broad immunity to incentivize activity that moderates such objectionable content but does not hold companies liable for what is left posted. State laws that are inconsistent with the scheme set out by Section 230 are expressly preempted.

The author argues that the bill walks this line carefully:

AB 587 does not impose liability based on the nature of content moderation decisions taken by social media platforms. Rather, the requirements of AB 587 are focused exclusively on disclosure of information relating to those practices, with liability imposed based on failure to disclose the specified information. By taking this transparency approach, AB 587 is thus unlikely to run afoul of the liability protections provided by Section 230, and would be far less susceptible to a preemption challenge than most attempts to regulate in this space.

In addition, any specific mandates to remove some subset of this broad swath of content could run afoul of the First Amendment. The United States Supreme Court has held that posting on social networking and/or social media sites constitutes communicative activity protected by the First Amendment.¹⁶ As a general rule, the government “may not suppress lawful speech as the means to suppress unlawful speech.”¹⁷ In addition, the First Amendment places restrictions on compelled speech. However, the case law generally affords a wide berth to laws that regulate commercial speech and that involve disclosure requirements that involve conveying factual information that has sound public policy justification, such as food labeling.¹⁸

Because this bill simply seeks transparency into what content moderation practices are being deployed and their outcomes, it likely does not run afoul of these laws. No specific content moderation is required or penalized. The information required to be disclosed can play a key role in informing future legislative action and public debate of these issues.

¹⁶ E.g., *Packingham v. North Carolina* (2017) 137 S.Ct. 1730, 1735-1736.

¹⁷ *Ashcroft v. Free Speech Coalition* (2002) 535 U.S. 234, 255; see also *United States v. Alvarez* (2012) 567 U.S. 709, 717 (Supreme Court “has rejected as ‘startling and dangerous’ a ‘free-floating test for First Amendment coverage...[based on] an ad hoc balancing of relative social costs and benefits’ ” [alterations in original]).

¹⁸ See *Central Hudson Gas & Elec. Corp. v. Public Serv. Comm’n* (1980) 447 U.S. 557; *Zauderer v. Office of Disciplinary Counsel of Supreme Court* (1985) 471 U.S. 626.

3. Defining social media

The bill defines “social media company” as a person or entity that owns or operates a public-facing internet-based service that generated at least \$100 million in gross revenue during the preceding calendar year, and that allows users in the state to do all of the following:

- construct a public or semipublic profile within a bounded system created by the service;
- populate a list of other users with whom an individual shares a connection within the system; and
- view and navigate a list of the individual’s connections and the connections made by other individuals within the system.

The bill specifically excludes from this definition persons or entities that exclusively own and operate an electronic mail service.

A number of companies and their representative organizations have argued that the definition unnecessarily sweeps up companies that most people would not consider “social media.” They request specific exemptions to carve them out of the definition in the bill.

There is evidence that an easy definition is elusive for many. Federal law does not provide a clear definition. For instance, the federal statute establishing the Social Media Working Group within the Department of Homeland Security does not even provide a definition. (6 USC 195d.) Other states that define social media in statute have used definitions similar to that laid out in the bill.¹⁹

In response to stakeholder concerns and at the request of the Committee to harmonize the relevant definitions in this bill and in AB 35 (Chau, 2021), the author has proposed a revised definition for social media company. One concern raised was that, while a company may operate a social media platform as part of its company, the entire company should not necessarily be subject to the provisions of this bill. The amendments make clear that the requirements of the bill apply to “social media platforms” which are owned or operated by a “social media company.” The definition stands apart from a clarifying section that follows, which simply indicates the entities that are not subject to the provisions of this particular bill. The proposed amendments are included at the end of this analysis.

¹⁹ See e.g., N.M. Stat. Ann. § 21-1-46; La. Rev. Stat. Ann. § 3:2462.

4. Concerns with the bill

Three major areas of concern that have been raised, primarily by a coalition in opposition to the bill, are the granularity of the data required to be reported, the reporting scheme itself, and the enforcement provisions. In response to the varied concerns of stakeholders, the author has proposed a series of amendments to the bill, which are included at the end of this analysis.

a. Narrowing what is included in the quarterly reports

The technology and business coalition in opposition, including TechNet and the California Chamber of Commerce, argue:

In seeking to increase transparency around content moderation practices, AB 587 requires companies to report to the Attorney General the guidelines, practices, and even training materials companies use to moderate their platforms. The recent amendments make it explicit that the bill is seeking “complete and detailed” information about content moderation practices, capabilities, and data regarding content moderation. This requirement would not only threaten the security of these practices but provides bad actors with roadmaps to get around our protections. We believe that while well intentioned, these requirements will ultimately allow scammers, spammers, and other bad actors to exploit our systems and moderators.

To avoid undermining the goals of the bill and the work our companies have already undertaken to combat harmful content, we suggest removing these requirements in order to prevent the disclosure of information that could be used against our platforms and our users.

Although opposition points out that many social media companies already post similar reports on their own websites, supporters believe more is necessary. Decode Democracy writes in support of requiring more:

While some platforms share limited information about their efforts, more transparency is needed to address concerns about the intent, enforcement, and impact of platform policies, and to provide policymakers and the general public with critical data and metrics regarding the scope and scale of online hate and disinformation. Greater transparency is needed so consumers can make informed choices about the impact of platform policies, (including on their children), and also to enable researchers, civil society leaders, and policymakers to determine the best way to address this threat to our democracy.

Ultimately, the goal of transparency in content moderation and terms of service is to understand what is being done in order to assess its effectiveness and to guide debate about what can be done to enhance that effectiveness. Arguably, this goal is undermined if every tactic to combat objectionable content, such as misinformation, hate speech, and outright criminal activity, was divulged in complete detail to those intending to subvert those very practices. Although the sophisticated actors that carry a majority of this out will likely already have much of this information, a better balance can be struck in the bill. In response, the author proposes a number of amendments to scale back the granularity of the data required to be placed in these reports. As seen by the language included at the end of this analysis, the author proposes removing the requirement to include items such as training manuals, a complete description of the company's rules and guidelines for content moderation, among other data points.

b. Reporting requirements

The coalition in opposition next argues the reporting requirements are too burdensome:

AB 587 requires businesses to report detailed metrics on a quarterly basis regarding not only the numerical scale of content moderation practices, but also details about how content is flagged and acted against. It would be nearly impossible to report this information quarterly due to the need to review, analyze, and adjudicate actioned content. Further, the sheer volume of content our companies review makes it similarly difficult and costly to implement these disclosures, particularly the number of times actioned items of content were viewed or shared. Producing this information quarterly is unworkable and unreasonable.

There is little justification to require a report to the Attorney General. Instead the bill should simply require our companies to post these reports on their website or platform.

The amendments discussed in the previous section will certainly mitigate the burden these reports impose. However, it is arguably unnecessarily onerous to require all of this information to be reported every three months. The author may wish to consider moving to longer reporting periods.

c. Enforcement

The bill provides that a violation of any of its provisions is actionable under the Unfair Competition Law. The requirement that a social media company post their terms of service, and include certain details, comes with a right to cure. Therefore, a company is not in violation of that section unless it fails to comply within 30 days of the Attorney General notifying them of their noncompliance.

The coalition in opposition asserts that the enforcement mechanisms are onerous and problematic, arguing against even allowing for injunctive relief. The coalition states:

AB 587 opens companies up to the threat of liability and government investigation for routine moderation practices. Companies should not be subject to civil penalties or injunctive relief for the filing of a report, especially as comprehensive as the ones contemplated by this bill. Such litigation will deter investment in content moderation and suppress ongoing efforts to protect users from harmful content online.

The bill simply seeks more transparency for better-informed decisionmaking on appropriate next steps. However, this crucial information will likely be hard to gather if there is no meaningful penalty for failing to comply. Despite the opposition's arguments, the provision providing for enforcement pursuant to the UCL is likely not even necessary, as violation of the bill's provisions would serve as a predicate offense for a UCL action. It should be noted that it is improbable that any private right of action would be afforded by this provision, given the nature of potential injuries. Therefore, enforcement is left to public prosecutors and counsel for local governments.

In addition, the right to cure before the Attorney General can enforce the requirement regarding posting terms of service is arguably unwarranted. Allowing noncompliance with the law until 30 days after the Attorney General investigates and determines there is noncompliance is not sound public policy and undermines meaningful enforcement. Given the repeated assertions by those in opposition that social media companies "already make their terms of service and community standards easily accessible on their websites," it does not seem an overly onerous requirement and noncompliance should be sufficient for the Attorney General to enforce within its discretion.

Ultimately, to ensure clarity in how enforcement is to be carried out, the author has proposed amendments, included below, that remove the right to cure and the UCL provisions. The proposed amendments insert a penalty scheme involving civil penalties for each violation for every day the violation continues.

SUPPORT

Anti-Defamation League (sponsor)
American Association of University Women, Camarillo Branch
Accountable Tech
American Jewish Committee - Los Angeles
The Arc and United Cerebral Palsy California Collaboration
Armenian Assembly of America
Armenian National Committee of America - Western Region
Asian Americans in Action
Bend the Arc: Jewish Action

Buen Vecino
California Asian Pacific American Bar Association
California Hawaii State Conference National Association for the Advancement of
Colored People
California League of United Latin American Citizens
Center for the Study of Hate & Extremism - California State University, San Bernardino
Common Sense
Decode Democracy
Esperanza Immigrant Rights Project, Catholic Charities of Los Angeles
The Greenlining Institute
Hindu American Foundation, Inc.
Islamic Networks Group
Israeli-American Civic Action Network
Japanese American Citizens League, Berkeley Chapter
Jewish Center for Justice
Jewish Public Affairs Committee
Korean American Bar Association of Northern California
Korean American Coalition - Los Angeles
League of United Latin American Citizens
National Association for the Advancement of Colored People, SV/SJ
Nailing It for America
National Council of Jewish Women, California
National Hispanic Media Coalition
Orange County Racial Justice Collaborative
Progressive Zionists of California
Rabbis and Cantor of Congregation or Ami
Sikh American Legal Defense and Education Fund (SALDEF)
Simon Wiesenthal Center, Inc.
Stonewall Democratic Club

OPPOSITION

California Chamber of Commerce
Chamber of Progress
Civil Justice Association of California
Consumer Technology Association
Internet Association
Internet Coalition
MPA - the Association of Magazine Media
Netchoice
TechNet

RELATED LEGISLATION

Pending Legislation:

SB 388 (Stern, 2021) requires a social media platform company, as defined, that, in combination with each subsidiary and affiliate of the service, has 25,000,000 or more unique monthly visitors or users for a majority of the preceding 12 months, to report to the Department of Justice by April 1, 2022, and annually thereafter, certain information relating to its efforts to prevent, mitigate the effects of, and remove potentially harmful content. SB 388 is pending before the Senate Judiciary Committee.

SB 435 (Cortese, 2021) provides, in relevant part, for a cause of action against an entity that publishes or republishes certain sexual content, as provided, and provides for civil penalties for every two hours flagged content is not taken down, as specified. SB 435 is pending before the Senate Judiciary Committee.

SB 746 (Skinner, 2021) requires businesses to disclose whether they use the personal information of consumers for political purposes, as defined, to consumers, upon request, and annually to the Attorney General or the California Privacy Protection Agency. This bill is currently on the Senate Floor.

AB 35 (Chau, 2021) requires a person that operates a social media platform to disclose whether or not the platform has a policy or mechanism in place to address the spread of misinformation. AB 35 is currently in this Committee and will be heard on the same day as this bill.

AB 1379 (Eduardo Garcia, 2021) requires an online platform that has 10,000,000 or more unique monthly United States visitors or users for a majority of months during the preceding 12 months that targets political advertising, as defined, to make available an application programming interface or other technical capability to enable qualified third parties to conduct independent analysis of bias and unlawful discriminatory impact of that targeted advertising. AB 1379 is pending before the Assembly Elections Committee.

AB 1114 (Gallagher, 2021) requires a social media company located in California to develop a policy or mechanism to address content or communications that constitute unprotected speech, including obscenity, incitement of imminent lawless action, and true threats, or that purport to state factual information that is demonstrably false. AB 1114 is pending before the Assembly Arts, Entertainment, Sports, Tourism, and Internet Media Committee.

AB 613 (Cristina Garcia, 2021) requires social media platforms, as defined, or users or advertisers posting on a social media platform, to place text or marking within or adjacent to retouched images that have been posted on the platform for promotional or

commercial purposes, and specify how that retouched image was altered. AB 613 is pending before the Assembly Privacy and Consumer Protection Committee.

Prior Legislation:

SB 890 (Pan, 2020) would have required social media companies to remove images and videos depicting crimes, as specified, and imposed civil penalties for failing to do so. SB 890 died in the Senate Judiciary Committee.

AB 2391 (Gallagher, 2020) would have prohibited social media sites from removing user-posted content on the basis of the political affiliation or viewpoint of that content, except where the social media site is, by its terms and conditions, limited to the promotion of only certain viewpoints and values and the removed content conflicts with those viewpoints or values. AB 2931 died in the Assembly Committee on Arts, Entertainment, Sports, Tourism, and Media.

AB 2442 (Chau, 2020) was substantially similar to this bill and would have required social media companies to disclose the existence, or lack thereof, of a misinformation policy, and imposed civil penalties for failing to do so. AB 2442 died in the Senate Judiciary Committee due to the COVID-19 pandemic.

AB 1316 (Gallagher, 2019) would have prohibited social media sites from removing user-posted content on the basis of the political affiliation or viewpoint of that content, except where the social media site is, by its terms and conditions, limited to the promotion of only certain viewpoints and values and the removed content conflicts with those viewpoints or values. AB 1316 was held on the floor of the Assembly and was re-introduced as AB 2931 (2020).

AB 288 (Cunningham, 2019) would have required a social networking service, at the request of a user, to permanently remove personally identifiable information and not sell the information to third parties, within a commercially reasonable time of the request. AB 288 died in the Assembly Committee on Privacy and Consumer Protection.

SB 1424 (Pan, 2018) would have established a privately funded advisory group to study the problem of the spread of false information through Internet-based social media platforms, and draft a model strategic plan for Internet-based social media platforms to use to mitigate this problem. SB 1424 was vetoed by Governor Brown, whose veto message stated that, as evidenced by the numerous studies by academic and policy groups on the spread of false information, the creation of a statutory advisory group to examine this issue is not necessary.

AB 3169 (Gallagher, 2018) would have prohibited social media sites from removing content on the basis of the political affiliation or viewpoint of the content, and prohibited internet search engines from removing or manipulating content from search

results on the basis of the political affiliation or viewpoint of the content. AB 3169 died in the Assembly Committee on Privacy and Protection.

SB 1361 (Corbett, 2010) would have prohibited social networking websites from displaying, to the public or other registered users, the home address or telephone number of a registered user of that site who is under 18 years of age, and imposed a civil penalty of up to \$10,000 for each willful and knowing violation of this prohibition. SB 1361 died in the Assembly Committee on Entertainment, Sports, Tourism, and Internet Media.

PRIOR VOTES:

Assembly Floor (Ayes 64, Noes 1)

Assembly Appropriations Committee (Ayes 13, Noes 0)

Assembly Judiciary Committee (Ayes 10, Noes 0)

Assembly Privacy and Consumer Protection Committee (Ayes 9, Noes 0)

AUTHOR'S PROPOSED AMENDMENTS

SEC. 2. Chapter 22.8 (commencing with Section 22675) is added to Division 8 of the Business and Professions Code, to read:

CHAPTER 22.8. Content Moderation Requirements for Internet Terms of Service

22675. For purposes of this chapter, the following definitions apply:

(a) "Actioned" means a social media company, due to a suspected or confirmed violation of the terms of service, has taken some form of action, including, but not limited to, removal, demonetization, deprioritization, or banning, against the relevant user or relevant item of content.

(b) "Content" means media, including, but not limited to, text, images, videos, and groups of users that are created, posted, shared, or otherwise interacted with by users on an internet-based service.

(c) ~~(1)~~ "Social media company" means a person or entity that owns or operates **#one or more social media platforms.**

(d) (1) "Social media platform" means an ~~public facing~~ internet-based service that ~~generated at least one hundred million dollars (\$100,000,000) in gross revenue during the preceding calendar year, and that~~ allows users ~~in the state~~ to do all of the following:

(A) Construct a public or semipublic profile within a bounded system created by the service.

(B) Populate a list of other users with whom an individual shares a connection within the system.

(C) View and navigate a list of ~~the individual's connections and the~~ connections made by other individuals within the system.

~~(2) "Social media company" does not include a person or entity that exclusively owns and operates an electronic mail service.~~

(d) "Terms of service" means a policy **or set of policies** adopted by a social media company that specifies, at least, the user behavior and activities that are permitted on the internet-based service owned or operated by the social media company, and the user behavior and activities that may subject the user or an item of content to being actioned. This may include, but is not limited to, a terms of service document or

agreement, rules or content moderation guidelines, community guidelines, acceptable uses, and other policies and established practices that outline these policies.

22676. (a) A social media company shall post ~~their~~ terms of service **for each social media platform owned or operated by that company** in a manner reasonably designed to inform all users of the **social media platform**~~internet-based service owned or operated by the social media company~~ of the existence and contents of the terms of service.

(b) The terms of service posted pursuant to subdivision (a) shall include all of the following:

(1) Contact information for the purpose of allowing users to ask the social media company questions about the terms of service.

(2) A description of the process that users must follow to flag content, groups, or other users that they believe violate the terms of service, and the social media company's commitments on response and resolution time.

(3) A list of potential actions the social media company may take against an item of content or a user, including, but not limited to, removal, demonetization, deprioritization, or banning.

(c) The terms of service posted pursuant to subdivision (a) shall be available in all **Medi-Cal threshold** languages, **as defined in subdivision (d) of Section 128552 of the Health and Safety Code**, in which the social media **platform**~~company~~ offers product features, including, but not limited to, menus and prompts.

~~(d) A social media company shall be in violation of this section if the social media company fails to comply with the provisions of this section within 30 days of being notified of noncompliance by the Attorney General.~~

22677. (a) On a quarterly basis, a social media company shall submit to the Attorney General a terms of service report, covering activity within the three months previous to the submission of the report. The terms of service report shall include, **for each social media platform owned or operated by the company, all of** the following:

(1) The current version of the terms of service of the social media **platform**~~company~~.

(2) If a social media company has filed its first quarterly report, a complete and detailed description of any changes to the terms of service since the last quarterly report.

(3) A statement of whether the current version of the terms of service defines each of the following categories of content, and, if so, the definitions of those categories, including any subcategories:

(A) Hate speech or racism.

(B) Extremism or radicalization.

(C) Disinformation or misinformation.

(D) Harassment.

(E) Foreign political interference.

(4) A ~~complete and~~ detailed description of content moderation practices used by the social media company **for that platform**, including, but not limited to, all of the following:

(A) Any existing policies intended to address the categories of content described in paragraph (3).

(B) ~~Any rules or guidelines regarding how~~ **How** a social media company's automated content moderation systems enforce terms of service **of the social media platform** and when these systems involve human review.

~~(C) Any training materials provided to human content moderators intended to educate them on the categories of content described in paragraph (3).~~

(D) How the social media company responds to user reports of violations of the terms of service.

(E) ~~Any rules, guidelines, product changes, and content moderator training materials that cover how~~ **How** the social media company would remove individual pieces of content, users, or groups that violate the terms of service, or take broader action against individual users or against groups of users that violate the terms of service.

(F) The languages in which the social media ~~company~~ **platform does not make terms of service available, but does offer** product features, including, but not limited to, menus and prompts ~~and the languages for which the social media company has terms of service.~~

(5) (A) Information on content that was flagged by the social media company as content belonging to any of the categories described in paragraph (3), including all of the following:

- (i) The total number of flagged items of content.
 - (ii) The total number of actioned items of content.
 - (iii) The total number of actioned items of content that resulted in action taken by the social media company against the user or group of users responsible for the content.
 - (iv) The total number of actioned items of content that were removed, demonetized, or deprioritized by the social media company.
 - (v) The number of times actioned items of content were viewed by users.
 - (vi) The number of times actioned items of content were shared, and the number of users that viewed the content before it was actioned.
 - (vii) The number of times users appealed social media company actions **taken on that platform** and the number of reversals of social media company actions on appeal disaggregated by each type of action.
- (B) All information required by subparagraph (A) shall be disaggregated into the following categories:
- (i) The category of content, including any relevant categories described in paragraph (3).
 - (ii) The type of content, including, but not limited to, posts, comments, messages, profiles of users, or groups of users.
 - (iii) The type of media of the content, including, but not limited to, text, images, and videos.
 - (iv) How the content was flagged, including, but not limited to, flagged by company employees or contractors, flagged by artificial intelligence software, flagged by community moderators, flagged by civil society partners, and flagged by users.
 - (v) How the content was actioned, including, but not limited to, actioned by company employees or contractors, actioned by artificial intelligence software, actioned by community moderators, actioned by civil society partners, and actioned by users.
- (b) A social media company shall submit its first terms of service report pursuant to subdivision (a) to the Attorney General no later than July 1, 2022.
- (c) The Attorney General shall post on its official website all terms of service reports submitted pursuant to this section.

22678. (a) A social media company that violates the provisions of this chapter shall be liable for a civil penalty not to exceed fifteen thousand dollars (\$15,000) per violation per day, and may be enjoined in any court of competent jurisdiction.

(b) Actions for relief pursuant to this chapter shall be prosecuted exclusively in a court of competent jurisdiction by the Attorney General or a district attorney or by a county counsel authorized by agreement with the district attorney in actions involving violation of a county ordinance, or by a city attorney of a city having a population in excess of 750,000, or by a city attorney in a city and county or, with the consent of the district attorney, by a city prosecutor in a city having a full-time city prosecutor in the name of the people of the State of California upon their own complaint or upon the complaint of a board, officer, person, corporation, or association.

(c) If an action pursuant to this section is brought by the Attorney General, one-half of the penalty collected shall be paid to the treasurer of the county in which the judgment was entered, and one-half to the General Fund. If the action is brought by a district attorney or county counsel, the penalty collected shall be paid to the treasurer of the county in which the judgment was entered. If the action is brought by a city attorney or city prosecutor, one-half of the penalty collected shall be paid to the treasurer of the city in which the judgment was entered, and one-half to the treasurer of the county in which the judgment was entered.

22679. (a) The duties and obligations imposed by this chapter are cumulative to any other duties or obligations imposed under local, state, or federal law and shall not be construed to relieve any party from any duties or obligations imposed under law.

(b) The remedies or penalties provided by this chapter are cumulative to each other and to any other remedies or penalties available under local, state, or federal law.

~~22678. A violation of this chapter is actionable under the Unfair Competition Law (Chapter 5 (commencing with Section 17200) of Part 2 of Division 7), in addition to any other applicable state or federal law.~~

22680. This chapter shall not apply to any of the following:

(a) A social media company that generated less than one hundred million dollars (\$100,000,000) in gross revenue during the preceding calendar year.

(b) A service that exclusively conveys electronic mail.

(c) A service that exclusively facilitates direct messaging between users.

(d) A section for user-generated comments on a digital news internet website that otherwise exclusively hosts content published by a person or entity described in subdivision (b) of Section 2 of Article I of the California Constitution.

(e) Consumer reviews of products or services on an internet website that serves the exclusive purpose of facilitating online commerce.

(f) An internet-based subscription streaming service that is offered to consumers for the exclusive purpose of transmitting licensed media, including audio or video files, in a continuous flow from the internet-based service to the end user, and does not host user-generated content.

(g) A service that operates for the exclusive purpose of cloud storage or shared document or file collaboration.

Exhibit 6

SENATE JUDICIARY COMMITTEE
Senator Thomas Umberg, Chair
2021-2022 Regular Session

AB 587 (Gabriel)
Version: June 23, 2022
Hearing Date: June 28, 2022
Fiscal: Yes
Urgency: No
CK

SUBJECT

Social media companies: terms of service

DIGEST

This bill requires social media companies, as defined, to post their terms of service and to submit quarterly reports to the Attorney General on their terms of service and content moderation policies and outcomes.

EXECUTIVE SUMMARY

In 2005, five percent of adults in the United States used social media. In just six years, that number jumped to half of all Americans. Today, over 70 percent of adults use at least one social media platform. Facebook alone is used by 69 percent of adults, and 70 percent of those adults say they use the platform on a daily basis.

Given the reach of social media platforms and the role they play in many people's lives, concerns have arisen over what content permeates these sites, entering the lives of the billions of users, and the effects that has on them and society as a whole. In particular, the sharpest calls for action focus on the rampant spread of misinformation, hate speech, and sexually explicit content. Social media companies' content moderation of a decade ago involved handfuls of individuals and user policies were minimal. These programs and policies have dramatically evolved over the years but the proliferation of objectionable content and "fake news" has led to calls for swifter and more aggressive action in response. However, there has also been backlash against perceived censorship in response to filtering of content and alleged "shadow banning."

This bill requires social media companies, as defined, to publicly post the terms of service, with certain required elements, for their social media platforms and to provide the Attorney General with a quarterly report on their content moderation procedures and outcomes.

This bill is sponsored by the Anti-Defamation League. It is supported by a variety of groups, including Common Sense and the Islamic Networks Group. It is opposed by various technology and business associations, including the California Chamber of Commerce and TechNet.

PROPOSED CHANGES TO THE LAW

Existing law:

- 1) Prohibits, through the United States Constitution, the enactment of any law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances. (U.S. Const. Amend. 1.)
- 2) Provides, through the California Constitution, for the right of every person to freely speak, write, and publish their sentiments on all subjects, being responsible for the abuse of this right. Existing law further provides that a law may not restrain or abridge liberty of speech or press. (Cal. Const., art. I, § 2(a).)
- 3) Provides, in federal law, that a provider or user of an interactive computer service shall not be treated as the publisher or speaker of any information provided by another information content provider. (47 U.S.C. § 230(c)(2).)
- 4) Provides that a provider or user of an interactive computer service shall not be held liable on account of:
 - a) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected; or
 - b) any action taken to enable or make available to information content providers or others the technical means to restrict access to such material. (47 U.S.C. § 230(c)(2).)
- 5) Defines “interactive computer service” as any information service, system, or access software provider that provides or enables computer access by multiple users to a computer server, including specifically a service or system that provides access to the Internet and such systems operated or services offered by libraries or educational institutions. (47 U.S.C. § 230(f)(2).)
- 6) Establishes the Unfair Competition Law (UCL) and defines “unfair competition” to mean and include any unlawful, unfair, or fraudulent business act or practice and unfair, deceptive, untrue, or misleading advertising and any act prohibited

by Chapter 1 (commencing with Section 17500) of Part 3 of Division 7 of the Business and Professions Code. (Bus. & Prof. Code § 17200 et seq.)

- 7) Provides that any person who engages, has engaged, or proposes to engage in unfair competition may be enjoined. Any person may pursue representative claims or relief on behalf of others only if the claimant meets specified standing requirements and complies with Section 382 of the Code of Civil Procedure, but these limitations do not apply to claims brought under this chapter by the Attorney General, or any district attorney, county counsel, city attorney, or city prosecutor in this state. (Bus. & Prof. Code § 17203.)
- 8) Requires actions for relief pursuant to the UCL be prosecuted exclusively in a court of competent jurisdiction and only by the following:
 - a) the Attorney General;
 - b) a district attorney;
 - c) a county counsel authorized by agreement with the district attorney in actions involving violation of a county ordinance;
 - d) a city attorney of a city having a population in excess of 750,000;
 - e) a city attorney in a city and county;
 - f) a city prosecutor in a city having a full-time city prosecutor in the name of the people of the State of California upon their own complaint or upon the complaint of a board, officer, person, corporation, or association with the consent of the district attorney; or
 - g) a person who has suffered injury in fact and has lost money or property as a result of the unfair competition. (Bus. & Prof. Code § 17204.)
- 9) Holds any person who engages, has engaged, or proposes to engage in unfair competition liable for a civil penalty not to exceed \$2,500 for each violation, which shall be assessed and recovered in a civil action brought by the Attorney General, or other public prosecutors. (Bus. & Prof. Code § 17206(a).)
- 10) Prohibits false or deceptive advertising to consumers about the nature of any property, product, or service, including false or misleading statements made in print, over the internet, or any other advertising method. (Bus. & Prof. Code § 17500.)
- 11) Defines libel as a false and unprivileged publication by writing, printing, or any other representation that exposes any person to hatred, contempt, ridicule, or obloquy, which causes that person to be shunned or avoided, or which has a tendency to injure that person in their occupation. (Civ. Code §§ 45, 47.)
- 12) Requires certain businesses to disclose the existence and details of specified policies, including:

- a) Operators of commercial websites or online services that collect personally identifiable information about individual consumers residing in California who use or visit the website must conspicuously post its privacy policy. (Bus. & Prof. Code § 22575.)
- b) Retailers and manufacturers doing business in this state and having annual worldwide gross receipts over \$100,000,000 must disclose online whether the business has a policy to combat human trafficking and, if so, certain details about that policy. (Civ. Code § 1714.43.)
- c) End-users of automated license plate recognition technology must post its usage and privacy policy on its website. (Civ. Code § 1798.90.53.)
- d) Campus bookstores at public postsecondary educational institutions must post in-store or online a disclosure of its retail pricing policy on new and used textbooks. (Educ. Code § 66406.7(f).)

This bill:

- 1) Requires a social media company to post terms of service for each social media platform owned or operated by the company in a manner reasonably designed to inform all users of the social media platform of the existence and contents of the terms of service. The terms of service shall include all of the following:
 - a) contact information for the purpose of allowing users to ask the social media company questions about the terms of service;
 - b) a description of the process that users must follow to flag content, groups, or other users that they believe violate the terms of service, and the social media company's commitments on response and resolution time; and
 - c) a list of potential actions the social media company may take against an item of content or a user, including, but not limited to, removal, demonetization, deprioritization, or banning.
- 2) Requires the terms of service to be available in all Medi-Cal threshold languages, as defined, in which the social media platform offers product features, including, but not limited to, menus and prompts.
- 3) Requires social media companies to submit a terms of service report, quarterly, with the first report due July 1, 2022, to the Attorney General, who must post it on their website. The terms of service report must include, for each social media platform owned or operated by the company, all of the following:
 - a) the current version of the terms of service of the social media platform;
 - b) if a social media company has filed its first quarterly report, a complete and detailed description of any changes to the terms of service since the last quarterly report;
 - c) a statement of whether the current version of the terms of service defines specified categories of content, and, if so, the definitions of those

- categories, including any subcategories. This includes hate speech, racism, extremism, harassment, disinformation, and foreign political interference;
- d) a detailed description of content moderation practices used by the social media company for that platform, including, but not limited to, all of the following:
- i. any policies intended to address the above categories of content;
 - ii. how automated content moderation systems enforce terms of service and when these systems involve human review;
 - iii. how the social media company responds to user reports of violations of the terms of service;
 - iv. how the social media company would remove individual pieces of content, users, or groups that violate the terms of service, or take broader action against individual users or against groups of users that violate the terms of service; and
 - v. the languages in which the social media platform does not make terms of service available, but does offer product features;
- e) information on content that was flagged by the social media company as content belonging to any of the above categories, including the total number of all of the following:
- i. flagged items of content;
 - ii. actioned items of content;
 - iii. actioned items of content that resulted in action taken by the social media company against the user or users responsible;
 - iv. actioned items of content that were removed, demonetized, or deprioritized by the social media company;
 - v. times actioned items of content were viewed by users;
 - vi. times actioned items of content were shared, and the number of users that viewed the content before it was actioned; and
 - vii. times users appealed social media company actions taken on that platform and the number of reversals on appeal disaggregated by each action; and
- f) all information required by (e) shall also be disaggregated into the category of content, the type of content, the type of media, and how the content was flagged and actioned.
- 4) Defines “social media company” as a person or entity that owns or operates one or more social media platforms, as defined. The bill exempts certain social media companies, including those making less than \$100,000,000 in gross revenue during the preceding year.
- 5) Defines “actioned” to mean a social media company, due to a suspected or confirmed violation of the terms of service, has taken some form of action, including, but not limited to, removal, demonetization, deprioritization, or banning, against the relevant user or relevant item of content.

- 6) Defines “terms of service” as a policy or set of policies adopted by a social media company that specifies, at least, the user behavior and activities that are permitted on the internet-based service owned or operated by the social media company, and the user behavior and activities that may subject the user or an item of content to being actioned. This may include, but is not limited to, a terms of service document or agreement, rules or content moderation guidelines, community guidelines, acceptable uses, and other policies and established practices that outline these policies.
- 7) Subject companies in violation to penalties of up to \$15,000 per violation per day to be sought by specified public prosecutors. The court is to assess the amount with consideration of whether the company made a reasonable, good faith attempt to comply. A social media company is in violation for each day it does any of the following:
 - a) fails to post terms of service;
 - b) fails to timely submit to the Attorney General the report required above;
or
 - c) materially omits or misrepresents required information in a submitted report.
- 8) Provides that the duties and obligations imposed are cumulative to any others imposed under local, state, or federal law. The remedies or penalties provided are also cumulative to each other and to any others available.

COMMENTS

1. Social media content

In recent years, the clamor for more robust content moderation on social media has reached a fever pitch. This includes calls to control disinformation or “fake news,” hate speech, political interference, and other online harassment.

The 2016 election was a major breaking point for many. Investigations uncovered attempted interference in the United States Presidential election through a social media “information warfare campaign designed to spread disinformation and societal division in the United States.”¹ The United States Senate Select Committee on Intelligence issued a report detailing how Russian operatives carried out their plan:

Masquerading as Americans, these operatives used targeted advertisements, intentionally falsified news articles, self-generated

¹ Select Committee on Intelligence, Russian Active Measures, Campaigns, and Interference in the 2016 U.S. Election, United States Senate, https://www.intelligence.senate.gov/sites/default/files/documents/Report_Volume2.pdf. All internet citations are current as of June 24, 2022.

content, and social media platform tools to interact with and attempt to deceive tens of millions of social media users in the United States. This campaign sought to polarize Americans on the basis of societal, ideological, and racial differences, provoked real world events, and was part of a foreign government's covert support of Russia's favored candidate in the U.S. presidential election.

This again became a threat in the 2020 election, with social media rife with misinformation such as the incorrect election date,² and then social media became a hotbed of misinformation about the results of the election.³ The author points to investigations that have found the violent insurrectionists that stormed the Capitol on January 6, 2021, were abetted and encouraged by posts on social media sites.⁴ In response to indications that social media provided a venue for those who overran and assaulted police officers, Facebook deflected blame, asserting that “these events were largely organized on platforms that don’t have our abilities to stop hate, don’t have our standards, and don’t have our transparency.”⁵ However, later indictments of those perpetrating the attack “made it clear just how large a part Facebook had played, both in spreading misinformation about election fraud to fuel anger among the Jan. 6 protesters, and in aiding the extremist militia’s communication ahead of the riots.”⁶

One area the author specifically focuses in on as motivation for the bill is the rise of hate speech online and the real world consequences. The author points to a recent study of over 500 million Twitter posts from 100 cities in the United States that found that “more targeted, discriminatory tweets posted in a city related to a higher number of hate crimes.”⁷

² Pam Fessler, *Robocalls, Rumors And Emails: Last-Minute Election Disinformation Floods Voters*, NPR (October 24, 2020), <https://www.npr.org/2020/10/24/927300432/robocalls-rumors-and-emails-last-minute-election-disinformation-floods-voters>.

³ Sheera Frenkel, *How Misinformation ‘Superspreaders’ Seed False Election Theories*, New York Times (November 23, 2020), <https://www.nytimes.com/2020/11/23/technology/election-misinformation-facebook-twitter.html>; Philip Bump, *The chain between Trump’s misinformation and violent anger remains unbroken*, Washington Post (May 12, 2021), <https://www.washingtonpost.com/politics/2021/05/12/chain-between-trumps-misinformation-violent-anger-remains-unbroken/>.

⁴ Ken Dilanian & Ben Collins, *There are hundreds of posts about plans to attack the Capitol. Why hasn't this evidence been used in court?* (April 20, 2021) NBC News, <https://www.nbcnews.com/politics/justice-department/we-found-hundreds-posts-about-plans-attack-capitol-why-aren-n1264291>.

⁵ Sheera Frenkel & Cecilia Kang, *Mark Zuckerberg and Sheryl Sandberg’s Partnership Did Not Survive Trump* (July 8, 2021) The New York Times, <https://www.nytimes.com/2021/07/08/business/mark-zuckerberg-sheryl-sandberg-facebook.html>.

⁶ *Ibid.*

⁷ Press Release, *Hate speech on Twitter predicts frequency of real-life hate crimes* (June 24, 2019) NYU Tandon School of Engineering, <https://engineering.nyu.edu/news/hate-speech-twitter-predicts-frequency-real-life-hate-crimes>.

Misinformation also poses a danger to public health: One study found that the more people rely on social media as their main news source, the more likely they are to believe misinformation about the COVID-19 pandemic.⁸ Another found that a mere 12 people are responsible for 65 percent of the false and misleading claims about COVID-19 vaccines on Facebook, Instagram, and Twitter.⁹ Misinformation hinders emergency responses to natural responses when social media posts contain incorrect or out-of-date information.¹⁰

The author frames the problem:

Over the past several years, there has been growing concern around the role of social media in promoting hate speech, disinformation, conspiracy theories, violent extremism, and severe political polarization. If properly managed, the ability for social media to amplify ideas and messages that would otherwise lack widespread exposure can give voice to otherwise marginalized populations and improve the public discourse, but the same capacity can feed the propagation of misinformation and dangerous rhetoric.

Writing in support, the Anti-Defamation League, the sponsor of this bill, further explains the context of the bill:

In recent years, there has been growing concern around the role of social media in promoting hate speech, disinformation, conspiracy theories, violent extremism, harassment, and severe political polarization. According to ADL's 2021 Online Hate and Harassment Survey, 41% of individuals experience online harassment and one in three of those individuals attribute at least some harassment to their identity. Identity-based harassment remains worrisome, affecting the ability of already marginalized communities to be safe in digital spaces.

Importantly, this hate and harassment isn't only taking place in the dark corners of the internet. 75% of ADL's 2021 Online Hate and Harassment Survey respondents who were harassed said at least some harassment

⁸ Yan Su, *It doesn't take a village to fall for misinformation: Social media use, discussion heterogeneity preference, worry of the virus, faith in scientists, and COVID-19-related misinformation belief* (May 2021) Telematics and Information, Vol. 58,

<https://www.sciencedirect.com/science/article/abs/pii/S0736585320302069?via%3Dihub>.

⁹ Shannon Bond, *Just 12 People Are Behind Most Vaccine Hoaxes On Social Media, Research Shows* (May 14, 2021) NPR, <https://www.npr.org/2021/05/13/996570855/disinformation-dozen-test-facebooks-tweeters-ability-to-curb-vaccine-hoaxes>.

¹⁰ United States Department of Homeland Security, *Countering False Information on Social Media in Disasters and Emergencies* (March 2018), https://www.dhs.gov/sites/default/files/publications/SMWG_Countering-False-Info-Social-Media-Disasters-Emergencies_Mar2018-508.pdf.

happened on Facebook – and many also attributed harassment to other mainstream social media platforms. And online extremism is also front and center: Facebook’s own researchers found that 64% of people who joined an extremist group on Facebook only did so because the company’s algorithm recommended it to them.

A recent Congressional Research Services Report discussed the issue of content moderation and specifically the spread of misinformation and the role that social media companies play in worsening the issue:

Two features of social media platforms—the user networks and the algorithmic filtering used to manage content—can contribute to the spread of misinformation. Users can build their own social networks, which affect the content that they see, including the types of misinformation they may be exposed to. Most social media operators use algorithms to sort and prioritize the content placed on their sites. These algorithms are generally built to increase user engagement, such as clicking links or commenting on posts. In particular, social media operators that rely on advertising placed next to user-generated content as their primary source of revenue have incentives to increase user engagement. These operators may be able to increase their revenue by serving more ads to users and potentially charging higher fees to advertisers. Thus, algorithms may amplify certain content, which can include misinformation, if it captures users’ attention.¹¹

The role that content moderation, or the lack of it, has in alleviating or exacerbating these issues has been a source of much debate. A policy paper published by the Shorenstein Center on Media, Politics, and Public Policy at the Harvard Kennedy School, *Countering Negative Externalities in Digital Platforms*, focuses on the costs associated with various internet platforms that are not absorbed by the companies themselves:

Today, in addition to the carcinogenic effects of chemical runoffs and first and second hand tobacco smoke, we have to contend with a new problem: the poisoning of our democratic system through foreign influence campaigns, intentional dissemination of misinformation, and incitements to violence inadvertently enabled by Facebook, YouTube and our other major digital platform companies.¹²

¹¹ Jason A. Gallo & Clare Y. Cho, *Social Media: Misinformation and Content Moderation Issues for Congress* (January 27, 2021) Congressional Research Service, <https://crsreports.congress.gov/product/pdf/R/R46662>.

¹² *Countering Negative Externalities in Digital Platforms* (October 7, 2019) Shorenstein Center on Media, Politics and Public Policy, <https://shorensteincenter.org/countering-negative-externalities-in-digital-platforms/>.

The paper asserts that these major platform companies “enable exceptionally malign activities” and “experience shows that the companies have not made sufficient investments to eliminate or reduce these negative externalities.”

As pointed out by recent Wall Street Journal reporting, the companies’ employees are aware of the dangers:

A Facebook Inc. team had a blunt message for senior executives. The company’s algorithms weren’t bringing people together. They were driving people apart.

“Our algorithms exploit the human brain’s attraction to divisiveness,” read a slide from a 2018 presentation. “If left unchecked,” it warned, Facebook would feed users “more and more divisive content in an effort to gain user attention & increase time on the platform.”

That presentation went to the heart of a question dogging Facebook almost since its founding: Does its platform aggravate polarization and tribal behavior?

The answer it found, in some cases, was yes.¹³

A recent New York Times article on leadership at Facebook elaborates:

To achieve its record-setting growth, the [Facebook] had continued building on its core technology, making business decisions based on how many hours of the day people spent on Facebook and how many times a day they returned. Facebook’s algorithms didn’t measure if the magnetic force pulling them back to Facebook was the habit of wishing a friend happy birthday, or a rabbit hole of conspiracies and misinformation.

Facebook’s problems were features, not bugs.¹⁴

Another paper recently released provides “Recommendations to the Biden Administration,” and is relevant to the considerations here:

The Administration should work with Congress to develop a system of financial incentives to encourage greater industry attention to the social

¹³ Jeff Horowitz & Deepa Seetharaman, *Facebook Executives Shut Down Efforts to Make the Site Less Divisive* (May 26, 2020) Wall Street Journal, <https://www.wsj.com/articles/facebook-knows-it-encourages-division-topexecutives-nixed-solutions-11590507499>.

¹⁴ Sheera Frenkel & Cecilia Kang, *Mark Zuckerberg and Sheryl Sandberg’s Partnership Did Not Survive Trump* (July 8, 2021) The New York Times, <https://www.nytimes.com/2021/07/08/business/mark-zuckerberg-sheryl-sandberg-facebook.html>.

costs, or “externalities,” imposed by social media platforms. A system of meaningful fines for violating industry standards of conduct regarding harmful content on the internet is one example. In addition, the Administration should promote greater transparency of the placement of digital advertising, the dominant source of social media revenue. This would create an incentive for social media companies to modify their algorithms and practices related to harmful content, which their advertisers generally seek to avoid.¹⁵

2. Content moderation, transparency, and the low-grade war on our cognitive security

There are a number of considerations when addressing how to approach the proliferation of these undesirable social media posts and the companies’ practices that fuel the flames. A number of methods of content moderation are being deployed and have evolved from simply blocking content or banning accounts to quarantining topics, removing posts from search results, barring recommendations, and down ranking posts in priority. However, there is a lack of transparency and understanding of exactly what companies are doing and why it does not seem to be enough. An article in the MIT Technology Review articulates the issues with content moderation behind the curtain:

As social media companies suspended accounts and labeled and deleted posts, many researchers, civil society organizations, and journalists scrambled to understand their decisions. The lack of transparency about those decisions and processes means that—for many—the election results end up with an asterisk this year, just as they did in 2016.

What actions did these companies take? How do their moderation teams work? What is the process for making decisions? Over the last few years, platform companies put together large task forces dedicated to removing election misinformation and labeling early declarations of victory. Sarah Roberts, a professor at UCLA, has written about the invisible labor of platform content moderators as a shadow industry, a labyrinth of contractors and complex rules which the public knows little about. Why don’t we know more?

In the post-election fog, social media has become the terrain for a low-grade war on our cognitive security, with misinformation campaigns and conspiracy theories proliferating. When the broadcast news business

¹⁵ Caroline Atkinson, et al., *Recommendations to the Biden Administration On Regulating Disinformation and Other Harmful Content on Social Media* (March 2021) Harvard Kennedy School & New York University Stern School of Business,
https://static1.squarespace.com/static/5b6df958f8370af3217d4178/t/6058a456ca24454a73370dc8/1616421974691/TechnologyRecommendations_2021final.pdf.

served the role of information gatekeeper, it was saddled with public interest obligations such as sharing timely, local, and relevant information. Social media companies have inherited a similar position in society, but they have not taken on those same responsibilities. This situation has loaded the cannons for claims of bias and censorship in how they moderated election-related content.

This bill seeks to increase transparency around what terms of service social media companies are setting out and how it ensures those terms are abided by. The goal is to learn more about the methods of content moderation and how successful they are. According to the author:

The line between providing an open forum for productive discourse and permitting the proliferation of hate speech and misinformation is a fine one, and depends largely on the structure and practices of the platform. However, these platforms rarely provide detailed insight into such practices, and into the relative effectiveness of different approaches. This, along with constraints imposed by existing federal law, has historically made policy-making in this space remarkably difficult. This bill seeks to provide critical transparency to both inform the public as to the policies and practices governing the content they post and engage with on social media, and to allow for comparative assessment of content moderation approaches to better equip both social media companies and policymakers to address these growing concerns.

A coalition of groups, including ADL, Equality California, NAACP, and Esperanza Immigrant Rights Project, emphasizes the need for the bill:

Despite the widespread nature of these concerns, efforts by social media companies to self-police such content have been widely criticized as opaque, arbitrary, biased, and inadequate. While some platforms share limited information about their efforts, the current lack of transparency has exacerbated concerns about the intent, enforcement, and impact of corporate policies, and deprived policymakers and the general public of critical data and metrics regarding the scope and scale of online hate and disinformation. Additional transparency is needed to allow consumers to make informed choices about the impact of these products (including the impact on their children) and so that researchers, civil society leaders, and policymakers can determine the best means to address this growing threat to our democracy.

AB 587 would address this troubling lack of transparency by requiring social media platforms to publicly disclose their policies and report key data and metrics around the enforcement of their policies. This disclosure

would be accomplished through quarterly public filings with the Attorney General.

This bill starts with a baseline requirement to have social media platforms post their terms of service. These policies must include information about how users can ask questions, how they can flag content or users in violation, and a list of potential actions that the company might take in response. To ensure meaningful access, the terms of service must be posted in a manner reasonably designed to inform all users of their existence and contents and available in all Medi-Cal threshold languages in which the social media platform offers product features.

The bill next requires an extremely detailed report to be compiled by these companies and submitted to the Attorney General on a quarterly basis. This report must include information on the terms of service, any changes made, and whether they define certain categories of content, including hate speech or racism; extremism or radicalization; disinformation or misinformation; harassment; and foreign political interference.

The bill also requires the report to contain a “detailed description of content moderation practices” used by the platform. There must also be outcome-focused information included. Platforms must report on the number of flagged items of content and the number of times the company took action in response. To understand the impact of the reported content, the report must detail the number of times this content was viewed and shared by users. The data must also include these details broken down by content category, the type of media, and other factors.

The bill leaves enforcement to public prosecutors who may seek injunctive relief and civil penalties of up to \$15,000 per violation per day, with the amount based on a consideration of whether the platform made a reasonable, good faith attempt to comply. A social media platform is subject to liability if it does the following:

- fails to post terms of service;
- fails to timely submit the quarterly report to the Attorney General; or
- materially omits or misrepresents required information in that report.

3. Legal Obstacles

As the author references above, all of this occurs within tight quarters due to federal statutory and constitutional law.

a. *Section 230*

Section 230 does not apply to the *users* of social media (or the internet generally), but rather applies to the *platforms themselves*. In the early 1990s, prior to the enactment of Section 230, two trial court orders – one in the United States District Court for the Southern District of New York, and New York state court – suggested that internet

platforms could be held liable for allegedly defamatory statements made by the platforms' users if the platforms engaged in any sort of content moderation (e.g., filtering out offensive material).¹⁶ In response, two federal legislators and members of the burgeoning internet industry crafted a law that would give internet platforms immunity from liability for users' statements, even if they might have reason to know that statements might be false, defamatory, or otherwise actionable.¹⁷ The result – Section 230 – was relatively uncontroversial at the time, in part because of the relative novelty of the internet and in part because Section 230 was incorporated into a much more controversial internet regulation scheme that was the subject of greater debate.¹⁸

The crux of Section 230 is laid out in two parts. The first provides that “[n]o provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.”¹⁹ The second provides a safe harbor for content moderation, by stating that no provider or user shall be held liable because of good-faith efforts to restrict access to material that is “obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected.”²⁰

Together, these two provisions give platforms immunity from any civil or criminal liability that could be incurred by user statements, while explicitly authorizing platforms to engage in their own content moderation without risking that immunity. Section 230 specifies that “[n]o cause of action may be brought and no liability may be imposed under any State law that is inconsistent with this section.”²¹ Courts have applied Section 230 in a vast range of cases to immunize internet platforms from “virtually all suits arising from third-party content.”²²

¹⁶ See *Cubby, Inc. v. Compuserve, Inc.* (S.D.N.Y. 1991) 776 F.Supp. 135, 141; *Stratton Oakmont v. Prodigy Servs. Co.* (N.Y. Sup. Ct., May 26, 1995) 1995 N.Y. Misc. LEXIS 229, *10-14. These opinions relied on case law developed in the context of other media, such as whether bookstores and libraries could be held liable for distributing defamatory material when they had no reason to know the material was defamatory. (See *Cubby, Inc.*, 776 F. Supp. at p. 139; *Smith v. California* (1959) 361 U.S. 147, 152-153.)

¹⁷ Kosseff, *The Twenty-Six Words That Created The Internet* (2019) pp. 57-65.

¹⁸ *Id.* at pp. 68-73. Section 230 was added to the Communications Decency Act of 1996 (title 5 of the Telecommunications Act of 1996, Pub. L. 104-104, 110 Stat. 56), which would have imposed criminal liability on internet platforms if they did not take steps to prevent minors from obtaining “obscene or indecent” material online. The Supreme Court invalidated the CDA, except for Section 230, on the basis that it violated the First Amendment. (See *Reno, supra*, 521 U.S. at p. 874.)

¹⁹ *Id.*, § 230(c)(1).

²⁰ *Id.*, § 230(c)(1) & (2).

²¹ *Id.*, § 230(e)(1) & (3).

²² Kosseff, *supra*, fn. 13, at pp. 94-95; see, e.g., *Doe v. MySpace Inc.* (5th Cir. 2008) 528 F.3d 413, 421-422; *Carfano v. Metrosplash.com, Inc.* (9th Cir. 2003) 339 F.3d 1119, 1125; *Zeran v. America Online, Inc.* (4th Cir. 1997) 129 F.3d 327, 333-334.

The author argues that the bill walks this line carefully:

AB 587 does not impose liability based on the nature of content moderation decisions taken by social media platforms. Rather, the requirements of AB 587 are focused exclusively on disclosure of information relating to those practices, with liability imposed based on failure to disclose the specified information. By taking this transparency approach, AB 587 is thus unlikely to run afoul of the liability protections provided by Section 230, and would be far less susceptible to a preemption challenge than most attempts to regulate in this space.

b. First Amendment

In addition, any specific mandates to remove some subset of this broad swath of content could run afoul of the First Amendment. The United States Supreme Court has held that posting on social networking and/or social media sites constitutes communicative activity protected by the First Amendment.²³ As a general rule, the government “may not suppress lawful speech as the means to suppress unlawful speech.”²⁴ In addition, the First Amendment places restrictions on compelled speech. However, the case law generally affords a wide berth to laws that regulate commercial speech and that involve disclosure requirements that involve conveying factual information that has sound public policy justification, such as food labeling.²⁵

Because this bill simply seeks transparency into what content moderation practices are being deployed and their outcomes, it likely does not run afoul of these laws. No specific content moderation is required or penalized. The information required to be disclosed can play a key role in informing future legislative action and public debate of these issues.

4. Defining social media platforms

With a multitude of bills currently moving through the legislative process that seek to regulate social media platforms, efforts have been made to harmonize the various definitions that exist. The author has agreed to amend in the following definition, which will be going into the various other bills. However, the Committee and authors will continue to engage with stakeholders to further refine the definition as necessary.

²³ E.g., *Packingham v. North Carolina* (2017) 137 S.Ct. 1730, 1735-1736.

²⁴ *Ashcroft v. Free Speech Coalition* (2002) 535 U.S. 234, 255; see also *United States v. Alvarez* (2012) 567 U.S. 709, 717 (Supreme Court “has rejected as ‘startling and dangerous’ a ‘free-floating test for First Amendment coverage...[based on] an ad hoc balancing of relative social costs and benefits’ ” [alterations in original]).

²⁵ See *Central Hudson Gas & Elec. Corp. v. Public Serv. Comm’n* (1980) 447 U.S. 557; *Zauderer v. Office of Disciplinary Counsel of Supreme Court* (1985) 471 U.S. 626.

Amendment

Replace definition of “content”:

(b) (1) “Content” means statements or comments made by users and media that are created, posted, shared, or otherwise interacted with by users on an internet-based service or application.

(2) “Content” does not include media put online exclusively for the purpose of cloud storage, transmitting documents, or file collaboration.

Replace definition of “social media platform”:

(d) “Social media platform” means a public or semipublic internet-based service or application that has users in California and that meets all of the following criteria:

(1) (A) A substantial function of the service or application is to connect users in order to allow users to interact socially with each other within the service or application.

(B) A service or application that provides email or direct messaging services shall not be considered to meet this criterion on the basis of that function alone.

(2) The service or application allows users to do all of the following:

(A) Construct a public or semipublic profile for purposes of signing into and using the service.

(B) Populate a list of other users with whom an individual shares a social connection within the system.

(C) Create or post content viewable by other users, including, but not limited to, on message boards, in chat rooms, or through a landing page or main feed that presents the user with content generated by other users.

(e) “Public or semipublic internet-based service or application” excludes a service or application used to facilitate communication within a business or enterprise among employees or affiliates of the business or enterprise, provided that access to the service or application is restricted to employees or affiliates of the business or enterprise using the service or application.

The author wishes to continue to limit the application of the bill to only platforms controlled by a business entity that generated at least \$100,000,000 in gross revenue during the preceding calendar year, so that exemption continues to apply. Most of the other exemptions are incorporated into the definition of social media platform, excluding email and direct messaging services from the social interaction element, for instance.

5. Concerns with the bill

A coalition, including TechNet, the Civil Justice Association, and the California Chamber of Commerce, objects to the granularity and depth of the data that is required to be provided and argues that disclosing their practices will allow bad actors to exploit these platforms:

AB 587 requires companies to publicly disclose more than just content moderation policies, which are already available to the public. The bill requires companies to report to the Attorney General sensitive information about how we implement policies, detect activity, train employees, and use technology to detect content in need of moderation. The language makes it explicit that the bill is seeking “detailed” information about content moderation practices, capabilities, and data regarding content moderation.

The coalition also argues the reporting requirements are too onerous:

AB 587 requires businesses to report detailed metrics on a quarterly basis regarding not only the numerical scale of content moderation practices, but also details about how content is flagged and acted against. It would be nearly impossible to report this information quarterly due to the need to review, analyze, and adjudicate actioned content. Further, the sheer volume of content our companies review makes it similarly difficult and costly to implement these disclosures, particularly the number of times actioned items of content were viewed or shared. Producing this information quarterly is unworkable and unreasonable.

Furthermore, there is little justification to require a report to the Attorney General. Instead, the bill should simply require our companies to post these reports on their website or platform. If the goal is increased transparency, it is unlikely that a consumer would ever look to the Attorney General’s website for information about a company’s terms of service or community guidelines. [. . .]

To address these concerns, we suggest amending the bill to require an annual report, to be posted in a manner reasonably designed to inform a platform’s users and to strike the requirements for specific information about company content moderation practices and training materials.

6. Support for this transparency measure

The Simon Wiesenthal Center writes in support:

AB 587 is a much-needed bill that addresses this public policy need. It will require social media platforms to publicly disclose their corporate policies and report key data and metrics around the enforcement of their policies. This disclosure would require corporate policies to be disclosed to the Attorney General on such issues as hate speech and racism, extremism and harassment. Corporations would be compelled to disclose how they enforce these policies and any changes they make to these policies or enforcement.

AB 587 is long overdue and the Simon Wiesenthal Center fully supports it.

Writing in support, the California Labor Federation points to evidence more needs to be done:

A GLAAD study in 2021 reported that 64 percent of LGBTQ social media users polled stated that they had experienced harassment and hate speech at a much higher rate than other identity groups on social media. Investigations have also shown that the violent riots at the U.S. Capitol in early January of 2021 were abetted and encouraged by posts on social media sites. Lack of accountability for social media platforms severely affects a wide range of marginalized communities—including, but not limited to, people of color, women, the LGBTQ+ community, Jewish and Muslim communities, and people seeking reproductive justice—who are disproportionately targeted by online hate and adversely affected by misinformation on social media.

The Los Angeles County Democratic Party explains its support:

Social media companies have behaved irresponsibly, allowing disinformation, misinformation and hate to add fuel to some of America's greatest challenges, like combating a global pandemic or grappling with longstanding social wounds around racism and misogyny, and they have also enabled the micro targeting of vulnerable individuals.

The requirements of AB 587, clear Terms of Service and their enforcement, regular reporting, and significant penalties for non-compliance, are necessary measures to bring some measure of corporate accountability to our increasingly virtual lives.

SUPPORT

Anti-Defamation League (sponsor)
Accountable Tech
Alameda County Democratic Party
American Academy of Pediatrics, California
American Association of University Women - California
American Association of University Women, Camarillo Branch
American Federation of State, County and Municipal Employees, AFL-CIO
American Jewish Committee - Los Angeles
American Jewish Committee - San Francisco
American Muslim & Multifaith Women's Empowerment Council
The Arc and United Cerebral Palsy California Collaboration
Armenian Assembly of America
Armenian National Committee of America - Western Region
Asian Americans in Action
Asian Law Alliance
Bend the Arc: Jewish Action
Buen Vecino
California Asian Pacific American Bar Association
California Federation of Teachers AFL-CIO
California Hawaii State Conference National Association for the Advancement of
Colored People
California Labor Federation, AFL-CIO
California League of United Latin American Citizens
California Nurses Association
California State Council of Service Employees International Union (SEIU California)
California Women's Law Center
Center for LGBTQ Economic Advancement & Research (CLEAR)
Center for the Study of Hate & Extremism - California State University, San Bernardino
College Democrats at UC Irvine
Common Sense
Consumer Reports Advocacy
Courage California
Davis College Democrats
Decode Democracy
Democratic Party of the San Fernando Valley
Democrats for Israel-Orange County
East Bay Young Democrats
Equality California
Esperanza Immigrant Rights Project, Catholic Charities of Los Angeles
The Greenlining Institute
Harvey Milk LGBTQ Democratic Club
Hindu American Foundation, Inc.

Islamic Networks Group
Islamic Networks Inc.
Israeli-American Civic Action Network
Japanese American Citizens League, Berkeley Chapter
Jewish Center for Justice
Jewish Family and Children's Services of San Francisco, the Peninsula, Marin and Sonoma Counties
Jewish Federation of Greater Los Angeles
Jewish Federation of The Sacramento Region and The Sacramento Jewish Community Relations Council
Jewish Public Affairs Committee
Korean American Bar Association of Northern California
Korean American Coalition - Los Angeles
League of United Latin American Citizens
Los Angeles County Democratic Party
Maplight
Miracle Mile Democratic Club
National Association for the Advancement of Colored People, SV/SJ
Nailing It for America
National Center for Lesbian Rights
National Council of Jewish Women, California
National Hispanic Media Coalition
Oakland Privacy
Orange County Racial Justice Collaborative
Pakistani-American Democratic Club of Orange County
Pilipino American Los Angeles Democrats (PALAD)
Progressive Zionists of California
Protect US
Rabbis and Cantor of Congregation or Ami
Sacramento County Young Democrats
Sacramento LGBT Community Center
San Fernando Valley Young Democrats
San Francisco Democratic Party
Santa Barbara Women's Political Committee
Sikh American Legal Defense and Education Fund (SALDEF)
Simon Wiesenthal Center, Inc.
The Source LGBT+ Center
Stonewall Democratic Club
United Food and Commercial Workers, Western States Council
United Nurses Associations of California/union of Health Care Professionals
Voices for Progress

OPPOSITION

California Chamber of Commerce
Chamber of Progress
Civil Justice Association of California
Computer and Communications Industry Association
Consumer Technology Association
Internet Coalition
MPA - the Association of Magazine Media
Netchoice
TechNet

RELATED LEGISLATION

Pending Legislation:

SB 1056 (Umberg, 2022) requires a social media platform, as defined, to clearly and conspicuously state whether it has a mechanism for reporting violent posts, as defined; and allows a person who is the target, or who believes they are the target, of a violent post to seek an injunction to have the violent post removed. This bill is currently in the Assembly Judiciary Committee.

AB 1628 (Ramos, 2022) requires online platforms to create and post a policy that includes policies regarding distribution of controlled substances and its prevention, reporting mechanisms, and resources. AB 1628 is currently pending before this Committee and will be heard on the same day as this bill.

AB 2273 (Wicks, 2022) establishes the California Age-Appropriate Design Code Act, placing a series of obligations and restriction on businesses that provide online services, products, or features likely to be accessed by a child. The bill tasks the California Privacy Protection Agency with establishing a taskforce to evaluate best practice and to adopt regulations. AB 2273 is currently pending before this Committee and will be heard on the same day as this bill.

AB 2408 (Cunningham, 2022) establishes a negligence cause of action for a platform's use of any design, feature, or affordance that causes a child user to become addicted to the platform. It also provides for heightened civil penalties in actions brought by public prosecutors. AB 2408 is currently pending before this Committee and will be heard on the same day as this bill.

AB 2879 (Low, 2022) requires social media platforms to implement a mechanism by which school administrators can report instances of cyberbullying, and to disclose specified data related to reported instances of cyberbullying and the platform's

response. AB 2879 is currently pending before this Committee and will be heard on the same day as this bill.

Prior Legislation:

SB 388 (Stern, 2021) would have required a social media platform company, as defined, that, in combination with each subsidiary and affiliate of the service, has 25,000,000 or more unique monthly visitors or users for a majority of the preceding 12 months, to report to the Department of Justice by April 1, 2022, and annually thereafter, certain information relating to its efforts to prevent, mitigate the effects of, and remove potentially harmful content. This bill died in the Senate Judiciary Committee.

SB 890 (Pan, 2020) would have required social media companies to remove images and videos depicting crimes, as specified, and imposed civil penalties for failing to do so. SB 890 died in the Senate Judiciary Committee.

AB 2391 (Gallagher, 2020) would have prohibited social media sites from removing user-posted content on the basis of the political affiliation or viewpoint of that content, except where the social media site is, by its terms and conditions, limited to the promotion of only certain viewpoints and values and the removed content conflicts with those viewpoints or values. AB 2931 died in the Assembly Committee on Arts, Entertainment, Sports, Tourism, and Media.

AB 2442 (Chau, 2020) was substantially similar to this bill and would have required social media companies to disclose the existence, or lack thereof, of a misinformation policy, and imposed civil penalties for failing to do so. AB 2442 died in the Senate Judiciary Committee due to the COVID-19 pandemic.

AB 1316 (Gallagher, 2019) would have prohibited social media sites from removing user-posted content on the basis of the political affiliation or viewpoint of that content, except where the social media site is, by its terms and conditions, limited to the promotion of only certain viewpoints and values and the removed content conflicts with those viewpoints or values. AB 1316 was held on the floor of the Assembly and was re-introduced as AB 2931 (2020).

AB 288 (Cunningham, 2019) would have required a social networking service, at the request of a user, to permanently remove personally identifiable information and not sell the information to third parties, within a commercially reasonable time of the request. AB 288 died in the Assembly Committee on Privacy and Consumer Protection.

SB 1424 (Pan, 2018) would have established a privately funded advisory group to study the problem of the spread of false information through Internet-based social media platforms, and draft a model strategic plan for Internet-based social media platforms to use to mitigate this problem. SB 1424 was vetoed by Governor Brown, whose veto

message stated that, as evidenced by the numerous studies by academic and policy groups on the spread of false information, the creation of a statutory advisory group to examine this issue is not necessary.

AB 3169 (Gallagher, 2018) would have prohibited social media sites from removing content on the basis of the political affiliation or viewpoint of the content, and prohibited internet search engines from removing or manipulating content from search results on the basis of the political affiliation or viewpoint of the content. AB 3169 died in the Assembly Committee on Privacy and Consumer Protection.

SB 1361 (Corbett, 2010) would have prohibited social networking websites from displaying, to the public or other registered users, the home address or telephone number of a registered user of that site who is under 18 years of age, and imposed a civil penalty of up to \$10,000 for each willful and knowing violation of this prohibition. SB 1361 died in the Assembly Committee on Entertainment, Sports, Tourism, and Internet Media.

PRIOR VOTES:

Assembly Floor (Ayes 64, Noes 1)

Assembly Appropriations Committee (Ayes 13, Noes 0)

Assembly Judiciary Committee (Ayes 10, Noes 0)

Assembly Privacy and Consumer Protection Committee (Ayes 9, Noes 0)

Exhibit 7

587

SENATE COMMITTEE ON APPROPRIATIONS

Senator Anthony Portantino, Chair

2021 - 2022 Regular Session

AB 587 (Gabriel) - Social media companies: terms of service

Version: June 30, 2022

Urgency: No

Hearing Date: August 1, 2022

Policy Vote: JUD. 9 - 0

Mandate: No

Consultant: Matthew Fleming

Bill Summary: AB 587 would require social media companies, as defined, to post their terms of service and to submit quarterly reports to the Attorney General on their terms of service and content moderation policies and outcomes.

Fiscal Impact:

- DOJ: The Department of Justice (DOJ) reports costs of \$414,000 in 2022-23 and \$711,000 annually thereafter in order to enforce the provisions of AB 587 and for IT resources to allow for submissions of terms of service (General Fund).
- Judicial Branch: Unknown cost pressures due to increased court workload (Special Fund – Trial Court Trust Fund, General Fund).

Background: In recent years, the clamor for more robust content moderation on social media has reached a fever pitch. This includes calls to control disinformation or “fake news,” hate speech, political interference, and other online harassment.

The 2016 election was a major breaking point for many. Investigations uncovered attempted interference in the United States Presidential election through a social media information warfare campaign designed to spread disinformation and societal division in the United States. The United States Senate Select Committee on Intelligence issued a report detailing how Russian operatives attempted to influence the U.S. presidential election.

This again became a threat in the 2020 election, with social media rife with misinformation such as the incorrect election date, and then social media became a hotbed of misinformation about the results of the election. The author points to investigations that have found the violent insurrectionists that stormed the Capitol on January 6, 2021, were abetted and encouraged by posts on social media sites.

Misinformation also poses a danger to public health: One study found that the more people rely on social media as their main news source, the more likely they are to believe misinformation about the COVID-19 pandemic. Another found that a mere 12 people are responsible for 65 percent of the false and misleading claims about COVID-19 vaccines on Facebook, Instagram, and Twitter. Misinformation hinders emergency responses to natural responses when social media posts contain incorrect or out-of-date information.

Proposed Law:

- Requires a social media company to post terms of service for each social media platform owned or operated by the company in a manner reasonably designed to inform all users of the social media platform of the existence and contents of the terms of service.
- Requires the terms of service to be available in all Medi-Cal threshold languages, as defined, in which the social media platform offers product features, including, but not limited to, menus and prompts.
- Requires social media companies to submit a terms of service report, quarterly, with the first report due July 1, 2022, to the Attorney General, who must post it on their website. Specifies the detail that is required to be included in the terms of service report.
- Subject companies in violation to penalties of up to \$15,000 per violation per day to be sought by specified public prosecutors.
- Provides that the duties and obligations imposed are cumulative to any others imposed under local, state, or federal law. The remedies or penalties provided are also cumulative to each other and to any others available.
- Defines terms for purposes of the bill's provisions.

Related Legislation:

- AB 2098 (Low, 2022) would designate COVID-19-related misinformation or disinformation spread by physicians as unprofessional conduct. AB 2098 is pending in this committee.
- AB 1628 (Ramos, 2022) requires online platforms to create and post a policy that includes policies regarding distribution of controlled substances and its prevention, reporting mechanisms, and resources. AB 1628 is pending on the Senate Floor
- AB 2273 (Wicks, 2022) establishes the California Age-Appropriate Design Code Act, placing a series of obligations and restriction on businesses that provide online services, products, or features likely to be accessed by a child. The bill tasks the California Privacy Protection Agency with establishing a taskforce to evaluate best practice and to adopt regulations. AB 2273 pending in this Committee.

Staff Comments:

- DOJ costs: The Consumer Protection Section within DOJ's Public Rights Division would be responsible for the enforcement of AB 587. To address the increase in workload, the Section anticipates requiring 1.0 Deputy Attorney General, 1.0 Associate Governmental Program Analyst, 1.0 Investigative Auditor, and the legal complement of 1.0 Legal Secretary beginning January 1, 2023 and ongoing.

- Court workload cost pressures: It is unknown how many additional civil actions would be brought by the AG as a result of the implementation of this bill. However, it generally costs about \$8,000 to operate a courtroom for one eight-hour day. Although courts are not funded on the basis of workload, increased pressure on the Trial Court Trust Fund and staff workload may create a need for increased funding for courts from the General Fund (GF) to perform existing duties. Numerous trial court operations are funded through the imposition and collection of criminal fines and fees. However, the Legislature has reduced and eliminated criminal fines and fees over the past five years. As a result, the 2022-23 budget includes an ongoing annual allocation of \$151.5 million and a one-time allocation of \$10.3 million from the General Fund in order to address declining revenue to the Trial Court Trust Fund.

-- END --

Exhibit 8

SENATE COMMITTEE ON APPROPRIATIONS

Senator Anthony Portantino, Chair
2021 - 2022 Regular Session

AB 587 (Gabriel) - Social media companies: terms of service

Version: June 30, 2022

Policy Vote: JUD. 9 - 0

Urgency: No

Mandate: No

Hearing Date: August 11, 2022

Consultant: Matthew Fleming

Bill Summary: AB 587 would require social media companies, as defined, to post their terms of service and to submit quarterly reports to the Attorney General on their terms of service and content moderation policies and outcomes.

******* ANALYSIS ADDENDUM – SUSPENSE FILE *******

**The following information is revised to reflect amendments
adopted by the committee on August 11, 2022**

Fiscal Impact:

- DOJ: The Department of Justice (DOJ) reports costs of \$414,000 in 2022-23 and \$711,000 annually thereafter in order to enforce the provisions of AB 587 and for IT resources to allow for submissions of terms of service (General Fund).
- Judicial Branch: Unknown cost pressures due to increased court workload (Special Fund – Trial Court Trust Fund, General Fund).

Author Amendments:

- Provide that social media companies must submit reports to the Attorney General on a semiannual, rather than a quarterly basis, as specified.
- Require reports to be submitted to the Attorney General electronically.
- Clarify that internet-based services or applications for which interactions between users are limited to direct messages, commercial interactions, consumer reviews of products, sellers, services, events, or places, or any combination such interactions, are exempt.
- Make other technical and clarifying amendments.

-- END --

Exhibit 9

Office of Senate Floor Analyses
(916) 651-1520 Fax: (916) 327-4478

THIRD READING

Bill No: AB 587
Author: Gabriel (D), et al.
Amended: 8/11/22 in Senate
Vote: 21

SENATE JUDICIARY COMMITTEE: 9-0, 6/28/22
AYES: Umberg, Caballero, Cortese, Durazo, Hertzberg, McGuire, Stern,
Wieckowski, Wiener
NO VOTE RECORDED: Borgeas, Jones

SENATE APPROPRIATIONS COMMITTEE: 5-0, 8/11/22
AYES: Portantino, Bradford, Laird, McGuire, Wieckowski
NO VOTE RECORDED: Bates, Jones

ASSEMBLY FLOOR: 64-1, 6/2/21 - See last page for vote

SUBJECT: Social media companies: terms of service

SOURCE: Anti-Defamation League

DIGEST: This bill requires social media companies, as defined, to post their terms of service and to submit reports to the Attorney General on their terms of service and content moderation policies and outcomes.

ANALYSIS:

Existing law:

- 1) Prohibits, through the United States Constitution, the enactment of any law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances. (U.S. Const. Amend. 1.)

- 2) Provides, through the California Constitution, for the right of every person to freely speak, write, and publish their sentiments on all subjects, being responsible for the abuse of this right. Existing law further provides that a law may not restrain or abridge liberty of speech or press. (Cal. Const., art. I, § 2(a).)
- 3) Provides, in federal law, that a provider or user of an interactive computer service shall not be treated as the publisher or speaker of any information provided by another information content provider. (47 U.S.C. § 230(c)(2).)
- 4) Provides that a provider or user of an interactive computer service shall not be held liable on account of:
 - a) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected; or
 - b) any action taken to enable or make available to information content providers or others the technical means to restrict access to such material. (47 U.S.C. § 230(c)(2).)
- 5) Establishes the Unfair Competition Law (UCL) and defines “unfair competition” to mean and include any unlawful, unfair, or fraudulent business act or practice and unfair, deceptive, untrue, or misleading advertising. (Bus. & Prof. Code § 17200 et seq.) Prohibits false or deceptive advertising to consumers about the nature of any property, product, or service, including false or misleading statements made in print, over the internet, or any other advertising method. (Bus. & Prof. Code § 17500.)
- 6) Requires certain businesses to disclose the existence and details of specified policies. (Bus. & Prof. Code § 22575; Civ. Code §§ 1714.43, 1798.90.53; Educ. Code § 66406.7(f).)

This bill:

- 1) Requires a social media company to post terms of service for each social media platform owned or operated by the company in a manner reasonably designed to inform all users of the social media platform of the existence and contents of the terms of service.

- 2) Requires the terms of service to be available in all Medi-Cal threshold languages, as defined, in which the social media platform offers product features, including, but not limited to, menus and prompts.
- 3) Requires social media companies to submit a terms of service report, on a semiannual basis to the Attorney General, who must make it available to the public in a searchable repository on its website.
- 4) Subject companies in violation to penalties of up to \$15,000 per violation per day to be sought by specified public prosecutors.

Comments

In 2005, five percent of adults in the United States used social media. In just six years, that number jumped to half of all Americans. Today, over 70 percent of adults use at least one social media platform. Facebook alone is used by 69 percent of adults, and 70 percent of those adults say they use the platform on a daily basis.

Given the reach of social media platforms and the role they play in many people's lives, concerns have arisen over what content permeates these sites, entering the lives of the billions of users, and the effects that has on them and society as a whole. In particular, the sharpest calls for action focus on the rampant spread of misinformation, hate speech, and sexually explicit content. Social media companies' content moderation of a decade ago involved handfuls of individuals and user policies were minimal. These programs and policies have dramatically evolved over the years but the proliferation of objectionable content and "fake news" has led to calls for swifter and more aggressive action in response. However, there has also been backlash against perceived censorship in response to filtering of content and alleged "shadow banning."

One area the author specifically focuses in on as motivation for the bill is the rise of hate speech online and the real world consequences. The author points to a recent study of over 500 million Twitter posts from 100 cities in the United States that found that "more targeted, discriminatory tweets posted in a city related to a higher number of hate crimes."¹

This bill seeks to increase transparency around what terms of service social media companies are setting out and how it ensures those terms are abided by. The goal is

¹ Press Release, *Hate speech on Twitter predicts frequency of real-life hate crimes* (June 24, 2019) NYU Tandon School of Engineering, <https://engineering.nyu.edu/news/hate-speech-twitter-predicts-frequency-real-life-hate-crimes>.

to learn more about the methods of content moderation and how successful they are.

According to the author:

The line between providing an open forum for productive discourse and permitting the proliferation of hate speech and misinformation is a fine one, and depends largely on the structure and practices of the platform. However, these platforms rarely provide detailed insight into such practices, and into the relative effectiveness of different approaches. This, along with constraints imposed by existing federal law, has historically made policy-making in this space remarkably difficult. This bill seeks to provide critical transparency to both inform the public as to the policies and practices governing the content they post and engage with on social media, and to allow for comparative assessment of content moderation approaches to better equip both social media companies and policymakers to address these growing concerns.

This bill starts with a baseline requirement to have social media platforms post their terms of service. These policies must include information about how users can ask questions, how they can flag content or users in violation, and a list of potential actions that the company might take in response. To ensure meaningful access, the terms of service must be posted in a manner reasonably designed to inform all users of their existence and contents and available in all Medi-Cal threshold languages in which the social media platform offers product features.

The bill next requires an extremely detailed report to be compiled by these companies and submitted to the Attorney General on a semiannual basis. The bill also requires the report to contain a “detailed description of content moderation practices” used by the platform. The bill leaves enforcement to public prosecutors.

This bill is sponsored by the Anti-Defamation League. It is supported by a variety of groups, including Common Sense and the Islamic Networks Group. It is opposed by various technology and business associations, including the California Chamber of Commerce and TechNet.

For a more thorough analysis of the bill, including a discussion of the relevant legal obstacles posed by the First Amendment and Section 230 of the

Communications Decency Act, please see the Senate Judiciary Committee analysis of the bill.

FISCAL EFFECT: Appropriation: No Fiscal Com.: Yes Local: No

According to the Senate Appropriations Committee:

- DOJ: The Department of Justice (DOJ) reports costs of \$414,000 in 2022-23 and \$711,000 annually thereafter in order to enforce the provisions of AB 587 and for IT resources to allow for submissions of terms of service (General Fund).
- Judicial Branch: Unknown cost pressures due to increased court workload (Special Fund – Trial Court Trust Fund, General Fund).

SUPPORT: (Verified 8/11/22)

Anti-Defamation League (source)

Accountable Tech

Alameda County Democratic Party

American Academy of Pediatrics, California

American Association of University Women - California

American Association of University Women, Camarillo Branch

American Federation of State, County and Municipal Employees, AFL-CIO

American Jewish Committee - Los Angeles

American Jewish Committee - San Francisco

American Muslim & Multifaith Women's Empowerment Council

The Arc and United Cerebral Palsy California Collaboration

Armenian Assembly of America

Armenian National Committee of America - Western Region

Asian Americans in Action

Asian Law Alliance

Bend the Arc: Jewish Action

Buen Vecino

California Asian Pacific American Bar Association

California Democratic Party

California Federation of Teachers AFL-CIO

California Hawaii State Conference National Association for the Advancement of Colored People

California Labor Federation, AFL-CIO

California League of United Latin American Citizens

California Nurses Association

California State Council of Service Employees International Union (SEIU California)
California Women's Law Center
Center for LGBTQ Economic Advancement & Research (CLEAR)
Center for the Study of Hate & Extremism - California State University, San Bernardino
City of San Luis Obispo
College Democrats at UC Irvine
Common Sense
Consumer Reports Advocacy
Courage California
Davis College Democrats
Decode Democracy
Democratic Party of the San Fernando Valley
Democrats for Israel-Orange County
East Bay Young Democrats
Equality California
Esperanza Immigrant Rights Project, Catholic Charities of Los Angeles
The Greenlining Institute
Harvey Milk LGBTQ Democratic Club
Hindu American Foundation, Inc.
Islamic Networks Group
Islamic Networks Inc.
Israeli-American Civic Action Network
Japanese American Citizens League, Berkeley Chapter
Jewish Center for Justice
Jewish Family and Children's Services of San Francisco, the Peninsula, Marin and Sonoma Counties
Jewish Federation of Greater Los Angeles
Jewish Federation of The Sacramento Region and The Sacramento Jewish Community Relations Council
Jewish Public Affairs Committee
Korean American Bar Association of Northern California
Korean American Coalition - Los Angeles
League of United Latin American Citizens
Los Angeles County Democratic Party
Maplight
Media Alliance
Miracle Mile Democratic Club
National Association for the Advancement of Colored People, SV/SJ

Nailing It for America
National Center for Lesbian Rights
National Council of Jewish Women, California
National Hispanic Media Coalition
Oakland Privacy
Orange County Racial Justice Collaborative
Pakistani-American Democratic Club of Orange County
Pilipino American Los Angeles Democrats
Planned Parenthood Affiliates of California
Progressive Zionists of California
ProtectUS
Rabbis and Cantor of Congregation or Ami
Sacramento County Young Democrats
Sacramento LGBT Community Center
San Diego City Attorney's Office
San Fernando Valley Young Democrats
San Francisco Democratic Party
Santa Barbara Women's Political Committee
Sikh American Legal Defense and Education Fund
Simon Wiesenthal Center, Inc.
The Source LGBT+ Center
Stonewall Democratic Club
United Food and Commercial Workers, Western States Council
United Nurses Associations of California/union of Health Care Professionals
Voices for Progress

OPPOSITION: (Verified 8/11/22)

California Chamber of Commerce
Chamber of Progress
Civil Justice Association of California
Computer and Communications Industry Association
Consumer Technology Association
Internet Coalition
MPA - the Association of Magazine Media
Netchoice
TechNet

ARGUMENTS IN SUPPORT: A coalition of groups, including ADL, Equality California, NAACP, and Esperanza Immigrant Rights Project, emphasizes the need for the bill:

“Despite the widespread nature of these concerns, efforts by social media companies to self-police such content have been widely criticized as opaque, arbitrary, biased, and inadequate. While some platforms share limited information about their efforts, the current lack of transparency has exacerbated concerns about the intent, enforcement, and impact of corporate policies, and deprived policymakers and the general public of critical data and metrics regarding the scope and scale of online hate and disinformation. Additional transparency is needed to allow consumers to make informed choices about the impact of these products (including the impact on their children) and so that researchers, civil society leaders, and policymakers can determine the best means to address this growing threat to our democracy.

AB 587 would address this troubling lack of transparency by requiring social media platforms to publicly disclose their policies and report key data and metrics around the enforcement of their policies. This disclosure would be accomplished through quarterly public filings with the Attorney General.”

ARGUMENTS IN OPPOSITION: A coalition, including TechNet, writes:

“AB 587 requires companies to publicly disclose more than just content moderation policies, which are already available to the public. The bill requires companies to report to the Attorney General sensitive information about how we implement policies, detect activity, train employees, and use technology to detect content in need of moderation. The language makes it explicit that the bill is seeking “detailed” information about content moderation practices, capabilities, and data regarding content moderation.”

ASSEMBLY FLOOR: 64-1, 6/2/21

AYES: Aguiar-Curry, Arambula, Bauer-Kahan, Bennett, Berman, Bloom, Boerner Horvath, Bryan, Burke, Calderon, Carrillo, Cervantes, Chau, Chiu, Cooley, Cooper, Cunningham, Daly, Davies, Frazier, Friedman, Gabriel, Gallagher, Cristina Garcia, Eduardo Garcia, Gipson, Lorena Gonzalez, Grayson, Holden, Irwin, Jones-Sawyer, Kalra, Lackey, Lee, Levine, Low, Maienschein, McCarty, Medina, Mullin, Muratsuchi, Nazarian, O'Donnell, Petrie-Norris, Quirk, Quirk-Silva, Ramos, Reyes, Luz Rivas, Robert Rivas, Rodriguez, Blanca Rubio, Salas, Santiago, Stone, Ting, Valladares, Villapudua, Waldron, Ward, Akilah Weber, Wicks, Wood, Rendon

NOES: Gray

NO VOTE RECORDED: Bigelow, Chen, Choi, Megan Dahle, Flora, Fong, Kiley, Mathis, Mayes, Nguyen, Patterson, Seyarto, Smith, Voepel

Prepared by: Christian Kurpiewski / JUD. / (916) 651-4113
8/15/22 13:15:10

****** END ******

Exhibit 10

Office of Senate Floor Analyses
(916) 651-1520 Fax: (916) 327-4478

THIRD READING

Bill No: AB 587
Author: Gabriel (D), et al.
Amended: 8/24/22 in Senate
Vote: 21

SENATE JUDICIARY COMMITTEE: 9-0, 6/28/22
AYES: Umberg, Caballero, Cortese, Durazo, Hertzberg, McGuire, Stern,
Wieckowski, Wiener
NO VOTE RECORDED: Borgeas, Jones

SENATE APPROPRIATIONS COMMITTEE: 5-0, 8/11/22
AYES: Portantino, Bradford, Laird, McGuire, Wieckowski
NO VOTE RECORDED: Bates, Jones

ASSEMBLY FLOOR: 64-1, 6/2/21 - See last page for vote

SUBJECT: Social media companies: terms of service

SOURCE: Anti-Defamation League

DIGEST: This bill requires social media companies, as defined, to post their terms of service and to submit reports to the Attorney General on their terms of service and content moderation policies and outcomes.

Senate Floor Amendments of 8/24/22 narrow the enforcement mechanism.

ANALYSIS:

Existing law:

- 1) Provides, through the California Constitution, for the right of every person to freely speak, write, and publish their sentiments on all subjects, being responsible for the abuse of this right. Existing law further provides that a law may not restrain or abridge liberty of speech or press. (Cal. Const., art. I, § 2(a).)

- 2) Provides, in federal law, that a provider or user of an interactive computer service shall not be treated as the publisher or speaker of any information provided by another information content provider. (47 U.S.C. § 230(c)(2).)
- 3) Limits the liability of a provider or user of an interactive computer service in connection with restricting access to certain materials. (47 U.S.C. § 230(c)(2).)
- 4) Establishes the Unfair Competition Law (UCL). (Bus. & Prof. Code § 17200 et seq.) Prohibits false or deceptive advertising to consumers about the nature of any property, product, or service, including false or misleading statements made in print, over the internet, or any other advertising method. (Bus. & Prof. Code § 17500.)
- 5) Requires certain businesses to disclose the existence and details of specified policies. (Bus. & Prof. Code § 22575; Civ. Code §§ 1714.43, 1798.90.53; Educ. Code § 66406.7(f).)

This bill:

- 1) Requires a social media company to post terms of service for each social media platform owned or operated by the company in a manner reasonably designed to inform all users of the social media platform of the existence and contents of the terms of service.
- 2) Requires the terms of service to be available in all Medi-Cal threshold languages, as defined, in which the social media platform offers product features, including, but not limited to, menus and prompts.
- 3) Requires social media companies to submit a terms of service report, on a semiannual basis to the Attorney General, who must make it available to the public in a searchable repository on its website.
- 4) Subject companies in violation to penalties of up to \$15,000 per violation per day to be sought by specified public prosecutors.

Comments

In 2005, five percent of adults in the United States used social media. In just six years, that number jumped to half of all Americans. Today, over 70 percent of

adults use at least one social media platform. Facebook alone is used by 69 percent of adults, and 70 percent of those adults say they use the platform on a daily basis.

Given the reach of social media platforms and the role they play in many people's lives, concerns have arisen over what content permeates these sites, entering the lives of the billions of users, and the effects that has on them and society as a whole. In particular, the sharpest calls for action focus on the rampant spread of misinformation, hate speech, and sexually explicit content. Social media companies' content moderation of a decade ago involved handfuls of individuals and user policies were minimal. These programs and policies have dramatically evolved over the years but the proliferation of objectionable content and "fake news" has led to calls for swifter and more aggressive action in response. However, there has also been backlash against perceived censorship in response to filtering of content and alleged "shadow banning."

One area the author specifically focuses in on as motivation for the bill is the rise of hate speech online and the real world consequences. The author points to a recent study of over 500 million Twitter posts from 100 cities in the United States that found that "more targeted, discriminatory tweets posted in a city related to a higher number of hate crimes."¹ This bill seeks to increase transparency around what terms of service social media companies are setting out and how it ensures those terms are abided by. The goal is to learn more about the methods of content moderation and how successful they are.

According to the author:

The line between providing an open forum for productive discourse and permitting the proliferation of hate speech and misinformation is a fine one, and depends largely on the structure and practices of the platform. However, these platforms rarely provide detailed insight into such practices, and into the relative effectiveness of different approaches. This, along with constraints imposed by existing federal law, has historically made policy-making in this space remarkably difficult. This bill seeks to provide critical transparency to both inform the public as to the policies and practices governing the content they post and engage with on social media, and to allow for comparative assessment of content moderation approaches to better equip both

¹ Press Release, *Hate speech on Twitter predicts frequency of real-life hate crimes* (June 24, 2019) NYU Tandon School of Engineering, <https://engineering.nyu.edu/news/hate-speech-twitter-predicts-frequency-real-life-hate-crimes>.

social media companies and policymakers to address these growing concerns.

This bill starts with a baseline requirement to have social media platforms post their terms of service. These policies must include information about how users can ask questions, how they can flag content or users in violation, and a list of potential actions that the company might take in response. To ensure meaningful access, the terms of service must be posted in a manner reasonably designed to inform all users of their existence and contents and available in all Medi-Cal threshold languages in which the social media platform offers product features. The bill next requires a detailed report to be compiled by these companies and submitted to the Attorney General on a semiannual basis. The bill also requires the report to contain a “detailed description of content moderation practices” used by the platform. The bill leaves enforcement to the Attorney General or specified city attorneys.

This bill is sponsored by the Anti-Defamation League. It is supported by a variety of groups, including Common Sense and the Islamic Networks Group. It is opposed by various technology and business associations, including the California Chamber of Commerce and TechNet.

For a more thorough analysis of the bill, including a discussion of the relevant legal obstacles posed by the First Amendment and Section 230 of the Communications Decency Act, please see the Senate Judiciary Committee analysis of the bill.

FISCAL EFFECT: Appropriation: No Fiscal Com.: Yes Local: No

According to the Senate Appropriations Committee:

- DOJ: The Department of Justice (DOJ) reports costs of \$414,000 in 2022-23 and \$711,000 annually thereafter in order to enforce the provisions of AB 587 and for IT resources to allow for submissions of terms of service (General Fund).
- Judicial Branch: Unknown cost pressures due to increased court workload (Special Fund – Trial Court Trust Fund, General Fund).

SUPPORT: (Verified 8/23/22)

Anti-Defamation League (source)
Accountable Tech

Alameda County Democratic Party
American Academy of Pediatrics, California
American Association of University Women - California
American Association of University Women, Camarillo Branch
American Federation of State, County and Municipal Employees, AFL-CIO
American Jewish Committee - Los Angeles
American Jewish Committee - San Francisco
American Muslim & Multifaith Women's Empowerment Council
The Arc and United Cerebral Palsy California Collaboration
Armenian Assembly of America
Armenian National Committee of America - Western Region
Asian Americans in Action
Asian Law Alliance
Bend the Arc: Jewish Action
Buen Vecino
California Asian Pacific American Bar Association
California Democratic Party
California Federation of Teachers AFL-CIO
California Hawaii State Conference National Association for the Advancement of
Colored People
California Labor Federation, AFL-CIO
California League of United Latin American Citizens
California Nurses Association
California State Council of Service Employees International Union (SEIU
California)
California Women's Law Center
Center for LGBTQ Economic Advancement & Research (CLEAR)
Center for the Study of Hate & Extremism - California State University, San
Bernardino
City of San Luis Obispo
College Democrats at UC Irvine
Common Sense
Consumer Reports Advocacy
Courage California
Davis College Democrats
Decode Democracy
Democratic Party of the San Fernando Valley
Democrats for Israel-Orange County
East Bay Young Democrats
Equality California

Esperanza Immigrant Rights Project, Catholic Charities of Los Angeles
The Greenlining Institute
Harvey Milk LGBTQ Democratic Club
Hindu American Foundation, Inc.
Islamic Networks Group
Islamic Networks Inc.
Israeli-American Civic Action Network
Japanese American Citizens League, Berkeley Chapter
Jewish Center for Justice
Jewish Family and Children's Services of San Francisco, the Peninsula, Marin and Sonoma Counties
Jewish Federation of Greater Los Angeles
Jewish Federation of The Sacramento Region and The Sacramento Jewish Community Relations Council
Jewish Public Affairs Committee
Korean American Bar Association of Northern California
Korean American Coalition - Los Angeles
League of United Latin American Citizens
Los Angeles County Democratic Party
Maplight
Media Alliance
Miracle Mile Democratic Club
National Association for the Advancement of Colored People, SV/SJ
Nailing It for America
National Center for Lesbian Rights
National Council of Jewish Women, California
National Hispanic Media Coalition
Oakland Privacy
Orange County Racial Justice Collaborative
Pakistani-American Democratic Club of Orange County
Pilipino American Los Angeles Democrats
Planned Parenthood Affiliates of California
Progressive Zionists of California
ProtectUS
Rabbis and Cantor of Congregation or Ami
Sacramento County Young Democrats
Sacramento LGBT Community Center
San Diego City Attorney's Office

San Fernando Valley Young Democrats
San Francisco Democratic Party
Santa Barbara Women's Political Committee
Sikh American Legal Defense and Education Fund
Simon Wiesenthal Center, Inc.
The Source LGBT+ Center
Stonewall Democratic Club
United Food and Commercial Workers, Western States Council
United Nurses Associations of California/union of Health Care Professionals
Voices for Progress

OPPOSITION: (Verified 8/23/22)

California Chamber of Commerce
Chamber of Progress
Civil Justice Association of California
Computer and Communications Industry Association
Consumer Technology Association
Internet Coalition
MPA - the Association of Magazine Media
Netchoice
TechNet

ARGUMENTS IN SUPPORT: A coalition of groups, including ADL, Equality California, NAACP, and Esperanza Immigrant Rights Project, emphasizes the need for the bill:

“Despite the widespread nature of these concerns, efforts by social media companies to self-police such content have been widely criticized as opaque, arbitrary, biased, and inadequate. While some platforms share limited information about their efforts, the current lack of transparency has exacerbated concerns about the intent, enforcement, and impact of corporate policies, and deprived policymakers and the general public of critical data and metrics regarding the scope and scale of online hate and disinformation. Additional transparency is needed to allow consumers to make informed choices about the impact of these products (including the impact on their children) and so that researchers, civil society leaders, and policymakers can determine the best means to address this growing threat to our democracy.

AB 587 would address this troubling lack of transparency by requiring social media platforms to publicly disclose their policies and report key data and

metrics around the enforcement of their policies. This disclosure would be accomplished through quarterly public filings with the Attorney General.”

ARGUMENTS IN OPPOSITION: A coalition, including TechNet, writes:

“AB 587 requires companies to publicly disclose more than just content moderation policies, which are already available to the public. The bill requires companies to report to the Attorney General sensitive information about how we implement policies, detect activity, train employees, and use technology to detect content in need of moderation. The language makes it explicit that the bill is seeking “detailed” information about content moderation practices, capabilities, and data regarding content moderation.”

ASSEMBLY FLOOR: 64-1, 6/2/21

AYES: Aguiar-Curry, Arambula, Bauer-Kahan, Bennett, Berman, Bloom, Boerner Horvath, Bryan, Burke, Calderon, Carrillo, Cervantes, Chau, Chiu, Cooley, Cooper, Cunningham, Daly, Davies, Frazier, Friedman, Gabriel, Gallagher, Cristina Garcia, Eduardo Garcia, Gipson, Lorena Gonzalez, Grayson, Holden, Irwin, Jones-Sawyer, Kalra, Lackey, Lee, Levine, Low, Maienschein, McCarty, Medina, Mullin, Muratsuchi, Nazarian, O'Donnell, Petrie-Norris, Quirk, Quirk-Silva, Ramos, Reyes, Luz Rivas, Robert Rivas, Rodriguez, Blanca Rubio, Salas, Santiago, Stone, Ting, Valladares, Villapudua, Waldron, Ward, Akilah Weber, Wicks, Wood, Rendon

NOES: Gray

NO VOTE RECORDED: Bigelow, Chen, Choi, Megan Dahle, Flora, Fong, Kiley, Mathis, Mayes, Nguyen, Patterson, Seyarto, Smith, Voepel

Prepared by: Christian Kurpiewski / JUD. / (916) 651-4113
8/26/22 15:32:05

**** **END** ****

Exhibit 11

Office of Senate Floor Analyses
(916) 651-1520 Fax: (916) 327-4478

THIRD READING

Bill No: AB 587
Author: Gabriel (D), et al.
Amended: 8/11/22 in Senate
Vote: 21

SENATE JUDICIARY COMMITTEE: 9-0, 6/28/22
AYES: Umberg, Caballero, Cortese, Durazo, Hertzberg, McGuire, Stern,
Wieckowski, Wiener
NO VOTE RECORDED: Borgeas, Jones

SENATE APPROPRIATIONS COMMITTEE: 5-0, 8/11/22
AYES: Portantino, Bradford, Laird, McGuire, Wieckowski
NO VOTE RECORDED: Bates, Jones

ASSEMBLY FLOOR: 64-1, 6/2/21 - See last page for vote

SUBJECT: Social media companies: terms of service

SOURCE: Anti-Defamation League

DIGEST: This bill requires social media companies, as defined, to post their terms of service and to submit reports to the Attorney General on their terms of service and content moderation policies and outcomes.

Senate Floor Amendments of 8/23/22 narrow the enforcement mechanism

ANALYSIS:

Existing law:

- 1) Provides, through the California Constitution, for the right of every person to freely speak, write, and publish their sentiments on all subjects, being responsible for the abuse of this right. Existing law further provides that a law may not restrain or abridge liberty of speech or press. (Cal. Const., art. I, § 2(a).)

- 2) Provides, in federal law, that a provider or user of an interactive computer service shall not be treated as the publisher or speaker of any information provided by another information content provider. (47 U.S.C. § 230(c)(2).)
- 3) Limits the liability of a provider or user of an interactive computer service in connection with restricting access to certain materials. (47 U.S.C. § 230(c)(2).)
- 4) Establishes the Unfair Competition Law (UCL). (Bus. & Prof. Code § 17200 et seq.) Prohibits false or deceptive advertising to consumers about the nature of any property, product, or service, including false or misleading statements made in print, over the internet, or any other advertising method. (Bus. & Prof. Code § 17500.)
- 5) Requires certain businesses to disclose the existence and details of specified policies. (Bus. & Prof. Code § 22575; Civ. Code §§ 1714.43, 1798.90.53; Educ. Code § 66406.7(f).)

This bill:

- 1) Requires a social media company to post terms of service for each social media platform owned or operated by the company in a manner reasonably designed to inform all users of the social media platform of the existence and contents of the terms of service.
- 2) Requires the terms of service to be available in all Medi-Cal threshold languages, as defined, in which the social media platform offers product features, including, but not limited to, menus and prompts.
- 3) Requires social media companies to submit a terms of service report, on a semiannual basis to the Attorney General, who must make it available to the public in a searchable repository on its website.
- 4) Subject companies in violation to penalties of up to \$15,000 per violation per day to be sought by specified public prosecutors.

Comments

In 2005, five percent of adults in the United States used social media. In just six years, that number jumped to half of all Americans. Today, over 70 percent of

adults use at least one social media platform. Facebook alone is used by 69 percent of adults, and 70 percent of those adults say they use the platform on a daily basis.

Given the reach of social media platforms and the role they play in many people's lives, concerns have arisen over what content permeates these sites, entering the lives of the billions of users, and the effects that has on them and society as a whole. In particular, the sharpest calls for action focus on the rampant spread of misinformation, hate speech, and sexually explicit content. Social media companies' content moderation of a decade ago involved handfuls of individuals and user policies were minimal. These programs and policies have dramatically evolved over the years but the proliferation of objectionable content and "fake news" has led to calls for swifter and more aggressive action in response. However, there has also been backlash against perceived censorship in response to filtering of content and alleged "shadow banning."

One area the author specifically focuses in on as motivation for the bill is the rise of hate speech online and the real world consequences. The author points to a recent study of over 500 million Twitter posts from 100 cities in the United States that found that "more targeted, discriminatory tweets posted in a city related to a higher number of hate crimes."¹ This bill seeks to increase transparency around what terms of service social media companies are setting out and how it ensures those terms are abided by. The goal is to learn more about the methods of content moderation and how successful they are.

According to the author:

The line between providing an open forum for productive discourse and permitting the proliferation of hate speech and misinformation is a fine one, and depends largely on the structure and practices of the platform. However, these platforms rarely provide detailed insight into such practices, and into the relative effectiveness of different approaches. This, along with constraints imposed by existing federal law, has historically made policy-making in this space remarkably difficult. This bill seeks to provide critical transparency to both inform the public as to the policies and practices governing the content they post and engage with on social media, and to allow for comparative assessment of content moderation approaches to better equip both

¹ Press Release, *Hate speech on Twitter predicts frequency of real-life hate crimes* (June 24, 2019) NYU Tandon School of Engineering, <https://engineering.nyu.edu/news/hate-speech-twitter-predicts-frequency-real-life-hate-crimes>.

social media companies and policymakers to address these growing concerns.

This bill starts with a baseline requirement to have social media platforms post their terms of service. These policies must include information about how users can ask questions, how they can flag content or users in violation, and a list of potential actions that the company might take in response. To ensure meaningful access, the terms of service must be posted in a manner reasonably designed to inform all users of their existence and contents and available in all Medi-Cal threshold languages in which the social media platform offers product features. The bill next requires a detailed report to be compiled by these companies and submitted to the Attorney General on a semiannual basis. The bill also requires the report to contain a “detailed description of content moderation practices” used by the platform. The bill leaves enforcement to the Attorney General or specified city attorneys.

This bill is sponsored by the Anti-Defamation League. It is supported by a variety of groups, including Common Sense and the Islamic Networks Group. It is opposed by various technology and business associations, including the California Chamber of Commerce and TechNet.

For a more thorough analysis of the bill, including a discussion of the relevant legal obstacles posed by the First Amendment and Section 230 of the Communications Decency Act, please see the Senate Judiciary Committee analysis of the bill.

FISCAL EFFECT: Appropriation: No Fiscal Com.: Yes Local: No

According to the Senate Appropriations Committee:

- DOJ: The Department of Justice (DOJ) reports costs of \$414,000 in 2022-23 and \$711,000 annually thereafter in order to enforce the provisions of AB 587 and for IT resources to allow for submissions of terms of service (General Fund).
- Judicial Branch: Unknown cost pressures due to increased court workload (Special Fund – Trial Court Trust Fund, General Fund).

SUPPORT: (Verified 8/23/22)

Anti-Defamation League (source)
Accountable Tech

Alameda County Democratic Party
American Academy of Pediatrics, California
American Association of University Women - California
American Association of University Women, Camarillo Branch
American Federation of State, County and Municipal Employees, AFL-CIO
American Jewish Committee - Los Angeles
American Jewish Committee - San Francisco
American Muslim & Multifaith Women's Empowerment Council
The Arc and United Cerebral Palsy California Collaboration
Armenian Assembly of America
Armenian National Committee of America - Western Region
Asian Americans in Action
Asian Law Alliance
Bend the Arc: Jewish Action
Buen Vecino
California Asian Pacific American Bar Association
California Democratic Party
California Federation of Teachers AFL-CIO
California Hawaii State Conference National Association for the Advancement of
Colored People
California Labor Federation, AFL-CIO
California League of United Latin American Citizens
California Nurses Association
California State Council of Service Employees International Union (SEIU
California)
California Women's Law Center
Center for LGBTQ Economic Advancement & Research (CLEAR)
Center for the Study of Hate & Extremism - California State University, San
Bernardino
City of San Luis Obispo
College Democrats at UC Irvine
Common Sense
Consumer Reports Advocacy
Courage California
Davis College Democrats
Decode Democracy
Democratic Party of the San Fernando Valley
Democrats for Israel-Orange County
East Bay Young Democrats
Equality California

Esperanza Immigrant Rights Project, Catholic Charities of Los Angeles
The Greenlining Institute
Harvey Milk LGBTQ Democratic Club
Hindu American Foundation, Inc.
Islamic Networks Group
Islamic Networks Inc.
Israeli-American Civic Action Network
Japanese American Citizens League, Berkeley Chapter
Jewish Center for Justice
Jewish Family and Children's Services of San Francisco, the Peninsula, Marin and Sonoma Counties
Jewish Federation of Greater Los Angeles
Jewish Federation of The Sacramento Region and The Sacramento Jewish Community Relations Council
Jewish Public Affairs Committee
Korean American Bar Association of Northern California
Korean American Coalition - Los Angeles
League of United Latin American Citizens
Los Angeles County Democratic Party
Maplight
Media Alliance
Miracle Mile Democratic Club
National Association for the Advancement of Colored People, SV/SJ
Nailing It for America
National Center for Lesbian Rights
National Council of Jewish Women, California
National Hispanic Media Coalition
Oakland Privacy
Orange County Racial Justice Collaborative
Pakistani-American Democratic Club of Orange County
Pilipino American Los Angeles Democrats
Planned Parenthood Affiliates of California
Progressive Zionists of California
ProtectUS
Rabbis and Cantor of Congregation or Ami
Sacramento County Young Democrats
Sacramento LGBT Community Center
San Diego City Attorney's Office

San Fernando Valley Young Democrats
San Francisco Democratic Party
Santa Barbara Women's Political Committee
Sikh American Legal Defense and Education Fund
Simon Wiesenthal Center, Inc.
The Source LGBT+ Center
Stonewall Democratic Club
United Food and Commercial Workers, Western States Council
United Nurses Associations of California/union of Health Care Professionals
Voices for Progress

OPPOSITION: (Verified 8/23/22)

California Chamber of Commerce
Chamber of Progress
Civil Justice Association of California
Computer and Communications Industry Association
Consumer Technology Association
Internet Coalition
MPA - the Association of Magazine Media
Netchoice
TechNet

ARGUMENTS IN SUPPORT: A coalition of groups, including ADL, Equality California, NAACP, and Esperanza Immigrant Rights Project, emphasizes the need for the bill:

“Despite the widespread nature of these concerns, efforts by social media companies to self-police such content have been widely criticized as opaque, arbitrary, biased, and inadequate. While some platforms share limited information about their efforts, the current lack of transparency has exacerbated concerns about the intent, enforcement, and impact of corporate policies, and deprived policymakers and the general public of critical data and metrics regarding the scope and scale of online hate and disinformation. Additional transparency is needed to allow consumers to make informed choices about the impact of these products (including the impact on their children) and so that researchers, civil society leaders, and policymakers can determine the best means to address this growing threat to our democracy.

AB 587 would address this troubling lack of transparency by requiring social media platforms to publicly disclose their policies and report key data and

metrics around the enforcement of their policies. This disclosure would be accomplished through quarterly public filings with the Attorney General.”

ARGUMENTS IN OPPOSITION: A coalition, including TechNet, writes:

“AB 587 requires companies to publicly disclose more than just content moderation policies, which are already available to the public. The bill requires companies to report to the Attorney General sensitive information about how we implement policies, detect activity, train employees, and use technology to detect content in need of moderation. The language makes it explicit that the bill is seeking “detailed” information about content moderation practices, capabilities, and data regarding content moderation.”

ASSEMBLY FLOOR: 64-1, 6/2/21

AYES: Aguiar-Curry, Arambula, Bauer-Kahan, Bennett, Berman, Bloom, Boerner Horvath, Bryan, Burke, Calderon, Carrillo, Cervantes, Chau, Chiu, Cooley, Cooper, Cunningham, Daly, Davies, Frazier, Friedman, Gabriel, Gallagher, Cristina Garcia, Eduardo Garcia, Gipson, Lorena Gonzalez, Grayson, Holden, Irwin, Jones-Sawyer, Kalra, Lackey, Lee, Levine, Low, Maienschein, McCarty, Medina, Mullin, Muratsuchi, Nazarian, O'Donnell, Petrie-Norris, Quirk, Quirk-Silva, Ramos, Reyes, Luz Rivas, Robert Rivas, Rodriguez, Blanca Rubio, Salas, Santiago, Stone, Ting, Valladares, Villapudua, Waldron, Ward, Akilah Weber, Wicks, Wood, Rendon

NOES: Gray

NO VOTE RECORDED: Bigelow, Chen, Choi, Megan Dahle, Flora, Fong, Kiley, Mathis, Mayes, Nguyen, Patterson, Seyarto, Smith, Voepel

Prepared by: Christian Kurpiewski / JUD. / (916) 651-4113
8/26/22 15:32:04

**** **END** ****

Exhibit 12

CONCURRENCE IN SENATE AMENDMENTS

AB 587 (Gabriel)

As Amended August 24, 2022

Majority vote

SUMMARY

This bill requires social media companies, as defined, to post their terms of service (ToS) in a manner reasonably designed to inform all users of specified policies and would require a social media company to submit semiannual reports, as specified, starting January 1, 2024, to the Attorney General (AG).

Senate Amendments

- 1) Reduce the frequency of required reports from quarterly to semiannually.
- 2) Limit the languages in which terms of service must be posted to only the Medi-Cal threshold languages.
- 3) Remove requirements to disclose specific rules, guidelines, product changes, or training materials, and remove the requirement that the detailed description of content moderation practices within a report be "complete."
- 4) Remove the 30-day right to cure violations.
- 5) Replace enforcement through the Unfair Competition Law with a specified enforcement mechanism that permits specified public attorneys to bring actions seeking injunctive relief in addition to a civil penalty not to exceed \$15,000 per violation per day.
- 6) Specify that a social media company shall be considered in violation of the provisions of the bill for each day the social media company fails to post terms of service, as specified, fails to timely submit to the AG a report as required, or materially omits or misrepresents required information in a report.
- 7) Limit enforcement of the bill's provisions to actions brought by the AG or by city attorneys of cities with populations over 750,000.
- 8) Offset reporting such that a submitted report need not be released until one full quarter after the period of time covered by the report, and specify the timing of these reports.
- 9) Clarify that the bill shall not be construed to apply to an internet-based service or application for which interactions between users are limited to direct messages, commercial transactions, consumer reviews of products, sellers, services, events, or places, or any combination thereof.
- 10) Consistent with the definition adopted for all pending bills pertaining to social media this legislative session, define "social media platform" to mean a public or semipublic internet-based service or application that has users in California and that meets both of the following criteria: a) a substantial function of the service or application is to connect users in order to allow users to interact socially with each other within the service or application, except as specified; and b) the service or application allows users to construct a public or semipublic

profile for purpose of signing into and using the service, populate a list of other users with whom an individual shares a social connection within the system, and create or post content viewable by others.

- 11) Specify exclusions from the provided definitions for "content" and "public or semipublic internet-based service or application", respectively, to exclude media put on a service or application exclusively for the purpose of cloud storage, transmitting files, or file collaboration, and services or applications used to facilitate communication within a business or enterprise exclusively among employees or affiliates of the business or enterprise.

COMMENTS

As online social media become increasingly central to the public discourse, the companies responsible for managing social media platforms are faced with a complex dilemma regarding content moderation, i.e., how the platforms determine what content warrants disciplinary action such as removal of the item or banning of the user. In broad terms, there is a general public consensus that certain types of content, such as child pornography, depictions of graphic violence, emotional abuse, and threats of physical harm, are undesirable, and should be mitigated on these platforms to the extent possible. Many other categories of information, however, such as hate speech, racism, extremism, misinformation, political interference, and harassment, are far more difficult to reliably define, and assignment of their boundaries is often fraught with political bias. In such cases, both action and inaction by these companies seems to be equally maligned: too much moderation and accusations of censorship and suppressed speech arise; too little, and the platform risks fostering a toxic, sometimes dangerous community.

AB 587 seeks to confront issues around social media content moderation practices by requiring the publication of ToS with specified information, and by requiring social media companies to submit semiannual reports containing information related to content moderation policies and data related to the application of those policies in practice. Though content moderation on social media is a notoriously difficult problem to tackle, AB 587 seemingly adopts a unique, data driven approach to progressing public policy in that space. Rather than placing specific content moderation requirements on companies, which in many cases raises constitutional issues, the bill instead provides for transparency and public accountability with respect to these practices, and establishes a timely, comprehensive dataset of untoward content on social media. This dataset can support research into the ever-changing social media ecosystem to help inform policies designed to root out its most problematic components while preserving its benefits for expression and connection.

Though granularity in this information can be useful for understanding the landscape and establishing transparency, however, opponents of the bill point out that too much granularity could put the platforms at risk. Indeed, in the past few years, the social media ecosystem has seen the emergence of sophisticated, sometimes state-sponsored actors seeking to exploit the design of their platforms toward nefarious ends. In this respect, it does not seem outlandish to presume that a large, detailed, public repository of information related to how content is moderated may increase sophistication of attempts to subvert content moderation systems. That said, in much the same way as policies for assessment and disclosure of security vulnerabilities is considered a best practice for cybersecurity, this same repository could enhance public scrutiny in a manner that would expose shortcomings in content moderation practices before they become catastrophic. Additionally, such information in aggregate from several platforms may facilitate

comparison and meta-analysis that can help establish best practices that, even if transparent, are nonetheless secure. Accordingly, on balance, it is difficult to determine whether extensive, detailed publication of moderation practices would increase or decrease the vulnerability of these platforms to exploitation by bad actors. Nonetheless, amendments taken in the Senate make minor changes to language enumerating required disclosures in order to clarify that disclosure of such highly granular information is not expected.

Senate amendments divorce the bill's provisions from enforcement through the Unfair Competition Law, instead specifying an enforcement mechanism that permits actions for relief to be prosecuted by a wide range of specified public attorneys. Specifically, if a social media company fails to post terms of service in accordance with the bill's requirements, fails to timely submit to the AG a required report, or materially omits or misrepresents required information in a report, that company shall be liable for a civil penalty not to exceed \$15,000 per violation per day, and may be enjoined in a court of competent jurisdiction.

Staff notes that the bill does not appear to require any particular actions on the part of the company other than: 1) posting terms of service in accordance with specified criteria; and 2) submitting semiannual reports containing specified information. As such, it would appear that violations of the bill would only occur if the company failed to perform one or both of these requirements, and that so long as the reports and ToS conform to the specifications, the actual content moderation itself is not subject to enforcement. It therefore does not appear likely that liability imposed by this bill would allow for lawsuits to be filed against platforms for the sufficiency of their moderation practices, arguably making the risk of preemption under Section 230 on these grounds minimal.

According to the Author

In recent years, there has been growing concern around the role of social media in promoting hate speech, disinformation, conspiracy theories, violent extremism, and severe political polarization. Twitter, along with other social media platforms, has been implicated as a venue for hate groups to safely grow. A recent study of Twitter posts from 100 [United States] cities found that the greater proportion of tweets related to race- and ethnicity-based discrimination in a given city, the more hate crimes were occurring in that city. Robert Bowers, accused of murdering 11 elderly worshipers at a Pennsylvania synagogue in 2018, had been active on Gab, a Twitter-like site used by white supremacists. Most recently, investigations have shown that the violent riots at the Capitol in early January of this year were abetted and encouraged by posts on social media sites.

AB 587 would require social media platforms to publicly disclose their corporate policies and report key data and metrics around the enforcement of their policies. This disclosure would be accomplished through semiannual public filings with the AG.

Arguments in Support

A coalition of civil, minority, and immigrant rights organizations including the Anti-Defamation League, Common Sense, and the California League of United Latin American Citizens argues:

[E]fforts by social media companies to self-police [problematic] content have been opaque, arbitrary, biased, and inadequate. While some platforms share limited information about their efforts, the current lack of transparency has exacerbated concerns about the intent, enforcement, and impact of corporate policies, and deprived policymakers and the general public of critical data and metrics regarding the scope and scale of online hate and

disinformation. Additional transparency is needed to allow consumers to make informed choices about the impact of these products (including on their children) and so that researchers, civil society leaders, and policymakers can determine the best means to address this growing threat to our democracy.

AB 587 would address this troubling lack of transparency by requiring social media platforms to publicly disclose their corporate policies and report key data and metrics around the enforcement of their policies.

Arguments in Opposition

A coalition of groups representing business interests including CalChamber, TechNet, and the Civil Justice Association of California argue in opposition unless amended:

AB 587 requires companies to publicly disclose more than just content moderation policies, which are already available to the public. The bill requires companies to report to the Attorney General sensitive information about how we implement policies, detect activity, train employees, and use technology to detect content in need of moderation. [...] This requirement would not only threaten the security of these practices but provides extremists, terrorist organizations, child predators, scammers, and serial abusers of our policies with roadmaps to get around our protections. [...] AB 587 will undermine the extensive work our companies have already undertaken to combat harmful content.

FISCAL COMMENTS

According to the Senate Appropriations Committee:

- 1) *DOJ*: The Department of Justice (DOJ) reports costs of \$414,000 in 2022-23 and \$711,000 annually thereafter to enforce the provisions of AB 587 and for IT resources to allow for submission of terms of service (General Fund)
- 2) *Judicial Branch*: Unknown cost pressures due to increased court workload (Special Fund - Trial Court Trust Fund, General Fund)
- 3)

VOTES:

ASM PRIVACY AND CONSUMER PROTECTION: 9-0-2

YES: Chau, Bauer-Kahan, Bennett, Carrillo, Cunningham, Gabriel, Irwin, Lee, Wicks

ABS, ABST OR NV: Kiley, Gallagher

ASM JUDICIARY: 10-0-1

YES: Stone, Gallagher, Chau, Chiu, Davies, Lorena Gonzalez, Holden, Kalra, Maienschein, Reyes

ABS, ABST OR NV: Kiley

ASM APPROPRIATIONS: 13-0-3

YES: Lorena Gonzalez, Calderon, Carrillo, Chau, Davies, Gabriel, Eduardo Garcia, Levine, Quirk, Robert Rivas, Akilah Weber, Holden, Luz Rivas

ABS, ABST OR NV: Bigelow, Megan Dahle, Fong

ASSEMBLY FLOOR: 64-1-14

YES: Aguiar-Curry, Arambula, Bauer-Kahan, Bennett, Berman, Bloom, Boerner Horvath, Bryan, Burke, Calderon, Carrillo, Cervantes, Chau, Chiu, Cooley, Cooper, Cunningham, Daly, Davies, Frazier, Friedman, Gabriel, Gallagher, Cristina Garcia, Eduardo Garcia, Gipson, Lorena Gonzalez, Grayson, Holden, Irwin, Jones-Sawyer, Kalra, Lackey, Lee, Levine, Low, Maienschein, McCarty, Medina, Mullin, Muratsuchi, Nazarian, O'Donnell, Petrie-Norris, Quirk, Quirk-Silva, Ramos, Reyes, Luz Rivas, Robert Rivas, Rodriguez, Blanca Rubio, Salas, Santiago, Stone, Ting, Valladares, Villapudua, Waldron, Ward, Akilah Weber, Wicks, Wood, Rendon

NO: Gray

ABS, ABST OR NV: Bigelow, Chen, Choi, Megan Dahle, Flora, Fong, Kiley, Mathis, Mayes, Nguyen, Patterson, Seyarto, Smith, Voepel

SENATE FLOOR: 33-3-4

YES: Allen, Archuleta, Atkins, Becker, Bradford, Caballero, Cortese, Dahle, Dodd, Durazo, Eggman, Glazer, Gonzalez, Hertzberg, Hueso, Hurtado, Kamlager, Laird, Leyva, Limón, McGuire, Min, Newman, Ochoa Bogh, Pan, Portantino, Roth, Rubio, Skinner, Stern, Umberg, Wieckowski, Wiener

NO: Melendez, Nielsen, Wilk

ABS, ABST OR NV: Bates, Borgeas, Grove, Jones

UPDATED

VERSION: August 24, 2022

CONSULTANT: Landon Klein / P. & C.P. / (916) 319-2200

FN: 0004499